

# Optimal micropatterns in transport networks

Alessio Brancolini<sup>\*†</sup>

Benedikt Wirth<sup>\*‡</sup>

30th November 2015

## Abstract

We consider two variational models for transport networks, an urban planning and a branched transport model, in which the degree of network complexity and ramification is governed by a small parameter  $\varepsilon > 0$ . Smaller  $\varepsilon$  leads to finer ramification patterns, and we analyse how optimal network patterns in a particular geometry behave as  $\varepsilon \rightarrow 0$  by proving an energy scaling law. This entails providing constructions of near-optimal networks as well as proving that no other construction can do better.

The motivation of this analysis is twofold. On the one hand, it provides a better understanding of the transport network models; for instance, it reveals qualitative differences in the ramification patterns of urban planning and branched transport. On the other hand, several examples of variational pattern analysis in the literature use an elegant technique based on relaxation and convex duality. Transport networks provide a relatively simple setting to explore variations and refinements of this technique, thereby increasing the scope of its applicability.

**Keywords:** micropatterns, energy scaling laws, optimal transport, optimal networks, branched transport, irrigation, urban planning, Wasserstein distance

*2010 MSC:* 49Q20, 49Q10, 90B10

## 1 Introduction

Pattern formation in physical experiments or in materials can sometimes be understood by means of so-called energy scaling laws. Typically, such an experiment can be described by a physical energy which is minimised subject to some boundary conditions. If the boundary conditions (or another global constraint) are incompatible with a homogeneous state, a pattern of rapid spatial oscillations between energetically favourable states can often be observed, which tries to recover compatibility with the boundary conditions. While an infinitely fine oscillation could achieve full compatibility, the physical energy usually prevents too fine oscillations via some regularising energy component so that the resulting pattern is a balance between the objectives of satisfying compatibility constraints and of keeping the regularising energy contribution small.

It is commonly impossible to find the truly optimal pattern, however, if the regularising energy term has a small weight parameter  $\varepsilon > 0$ , one can instead try to prove how the minimum energy scales in  $\varepsilon$ , a so-called energy scaling law. This involves proving a lower and an upper bound for the energy as a function of  $\varepsilon$ , where both bounds differ at most by a multiplicative constant. The benefit is that any pattern satisfying the upper bound must have optimal energy up to a constant factor (and thus indicates how near-optimal patterns look like), since the truly optimal energy still lies above the lower bound. The proof of the upper bound is by constructing an appropriate pattern with the desired energy scaling, while the proof of the lower bound, which shows that no construction can do better, is ansatz-free and thus more complicated (however, in some cases a particular simple technique can be employed, variations of which are explored in this article).

---

<sup>\*</sup>Institute for Numerical and Applied Mathematics, University of Münster, Einsteinstraße 62, D-48149 Münster, Germany

<sup>†</sup>Email address: [alessio.brancolini@uni-muenster.de](mailto:alessio.brancolini@uni-muenster.de)

<sup>‡</sup>Email address: [benedikt.wirth@uni-muenster.de](mailto:benedikt.wirth@uni-muenster.de)

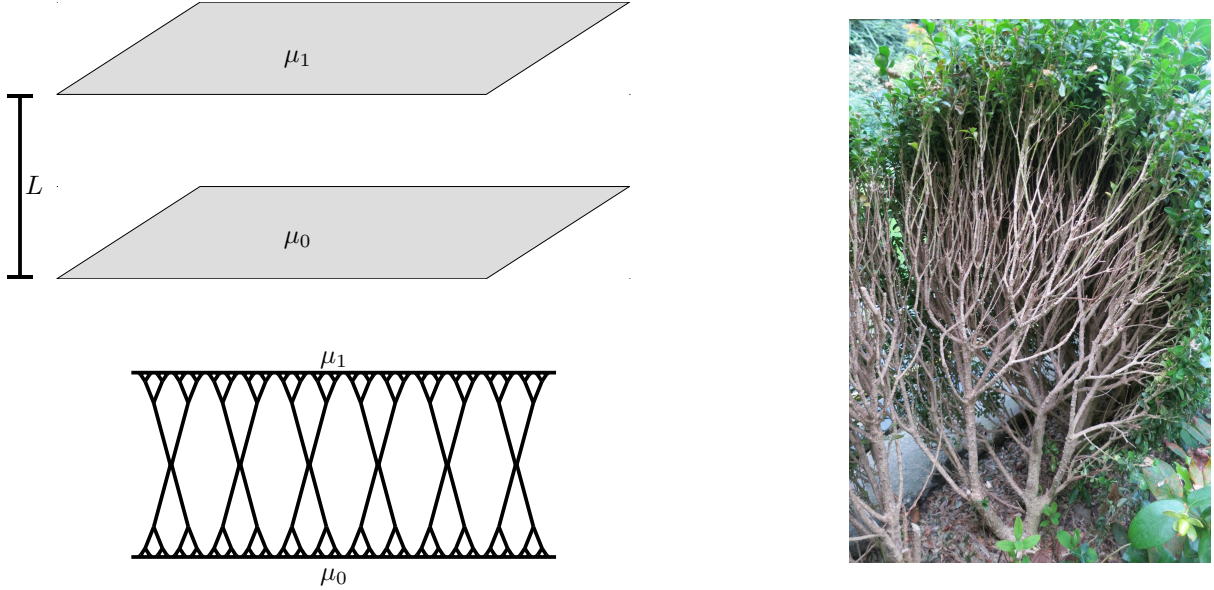


Figure 1: Top left: Sketch of the considered geometry, two measures  $\mu_0$  and  $\mu_1$  supported on codimension-one hypersquares at distance  $L$ . Bottom left: Exemplary transport network in two dimensions, transporting mass from  $\mu_0$  to  $\mu_1$ . Right: Photograph of a box tree exhibiting a branched network structure and a comparatively thin leaf canopy that may almost be thought of as a lower-dimensional manifold.

Energy scaling laws have been derived for various systems such as martensite–austenite transformations [KM92, KM94, KKO13, CC15, BG15], micromagnetics [CKO99], intermediate states in type-I superconductors [CKO04, CCKO08], membrane folding and blistering [BK14, BK15], or epitaxial growth [GZ14]. Energy scaling laws may not only be used to explain observed physical patterns, but they can also be applied to design problems [KW14, KW15], for instance, to find efficient engineering structures or to better understand why biological evolution has led to particular structures. This article represents a further contribution to the application of these techniques to design problems.

The setting we shall look at is a type of network optimisation to be explained in more detail in the next two sections. We consider two uniform measures  $\mu_0, \mu_1$  in  $\mathbb{R}^n$ , of same mass and supported on hyper-squares of codimension one, and we ask how the optimal transport network looks like to transport the mass in  $\mu_0$  to  $\mu_1$ . In more detail, let  $A \subset \mathbb{R}^{n-1}$  be a hyper-square and consider

$$\mu_0 = m\mathcal{H}^{n-1}\llcorner(A \times \{0\}), \quad \mu_1 = m\mathcal{H}^{n-1}\llcorner(A \times \{L\})$$

with  $m > 0$  and  $\mathcal{H}^{n-1}$  the  $(n-1)$ -dimensional Hausdorff measure (see Figure 1). Now  $\mu_0$  represents an initial and  $\mu_1$  the final distribution of particles, and we seek the optimal, locally one-dimensional transport network to move all particles. Optimality is with respect to a certain objective energy (in this article we consider a branched transport and an urban planning energy) which effectively encodes that a lumped transportation of many particles together is cheaper than transporting each particle individually. As a result, branched network structures as in Figure 1 are favoured, which balance the preference for lumping many particles together versus the restriction to evenly collect or distribute the particles on  $\mu_0$  and  $\mu_1$ , respectively. The degree of cost savings through lumped transportation will be described by a parameter  $\varepsilon > 0$ , and for  $\varepsilon > 0$  small, very fine network patterns will be optimal. We will analyse the (near-)optimal patterns by proving corresponding energy scaling laws in  $\varepsilon$ .

The motivation of this article and the goals pursued are multiple ones:

- The network models we examine and their generalisations have been used to model—among others—networks in public transport, pipeline systems, river geometry, and biology such as blood vessel systems or water transport paths in plants (see the extensive list of references in the monographs [BCM09, BPSS09]). Due to the strong interest in these models it is worthwhile to better understand their behaviour. Even though we consider a very special geometry for  $\mu_0$  and  $\mu_1$ , we will later argue that the same phenomena will be observed for geometry variations such as varying

the shape of the set  $A$  or replacing  $\mu_0$  by a delta distribution. The measure support on a set of codimension one may be seen as an approximation of cases where the extent of the initial or final mass distribution in the dominant transport direction is considerably smaller than its transversal extent as well as the typical transport length. Thinking of water transport to plant leaves, for instance, a single leaf or the leaf canopy may often be abstracted as a two-dimensional manifold embedded in three-dimensional space (Figure 1).

- Different network models have different mechanisms to give preference to branched network structures. The question arises whether the models exhibit more or less the same behaviour and thus are phenomenologically equivalent or whether one model can be told from the other just based on an observation of the network structure. We will see that the urban planning and the branched transport model, which we consider, share very similar features in three dimensions, but differ qualitatively in higher and lower dimensions. Indeed, while the energy distribution in branched transport is always rather evenly spread across all branching structures except the ones at the boundary, the major energy in urban planning is contributed by the coarsest structures in two dimensions and by the finest structures in dimensions four and higher.
- The (locally) one-dimensional networks that we consider may be viewed as approximation of transport channels that are very thin yet have a positive width. In that sense the models treated here can also be used to analyse structures for the transport of a heat or a magnetic flux as examined in [CCKO08, KW14]. In mathematical terms, it is expected (and currently investigated [Con]) that for smaller and smaller fluxes, the physical objective energy for conducting a heat or magnetic flux  $\Gamma$ -converges in the right scaling against the optimal network energy.
- The derivation of energy scaling laws is usually performed on a case to case basis, and no generally applicable method or theory has been developed so far. Instead, the field is still in the state of building and refining technique. In this article we explore how far one can advance based on a very simple and transparent technique originally introduced by Kohn and Müller [KM94]. The method mainly relies on convex duality, and appropriate variations of it indeed suffice to almost comprehensively understand energy scaling in optimal transport networks.

The remainder of the introduction provides some general notation and a reference list of symbols to be used in the network models. Section 2 then introduces the network models considered in this article, including references to the relevant literature, as well as the statement of the corresponding energy scaling laws Theorems 2.4.1 and 2.4.2, the main results of this article. Sections 3 and 4 contain the proofs of the upper bounds and the lower bounds, respectively, where both network models are treated separately. We close with a discussion in Section 5.

## 1.1 Preliminaries: notation and useful notions

Here we fix some frequently used basic notation.

- **Lebesgue measure.**  $\mathcal{L}^n$  denotes the  $n$ -dimensional *Lebesgue measure*.
- **Hausdorff measure.**  $\mathcal{H}^r$  denotes the  $r$ -dimensional *Hausdorff measure*.
- **Non-negative finite Borel measures.**  $\text{fbm}(X)$  denotes the set of *non-negative finite Borel measures* on a Borel set  $X \subset \mathbb{R}^n$ . Notice that these measures are countably additive and also regular by [Rud87, Thm. 2.18]. The corresponding total variation norm is denoted by  $\|\cdot\|_{\text{fbm}}$ .
- **(Signed or vector-valued) regular countably additive measures.**  $\text{rca}(X)$  denotes the set of *(signed or vector-valued) regular countably additive measures* on a Borel set  $X \subset \mathbb{R}^n$ . The corresponding total variation norm is denoted by  $\|\cdot\|_{\text{rca}}$ .
- **Weak-\* convergence.** The weak-\* convergence on  $\text{fbm}(\mathbb{R}^n)$  or  $\text{rca}(\mathbb{R}^n)$  is indicated by  $\xrightarrow{*}$ .
- **Restriction of a measure to a set.** Let  $(X, \mathcal{M}, \mu)$  be a measure space and  $Y \in \mathcal{M}$ . The measure  $\mu \llcorner Y$  is the measure defined by

$$\mu \llcorner Y(A) = \mu(A \cap Y).$$

- **Pushforward of a measure.** For a measure space  $(X, \mathcal{M}, \mu)$ , a measurable space  $(Y, \mathcal{N})$ , and a measurable map  $T : X \rightarrow Y$ , the *pushforward* of  $\mu$  under  $T$  is the measure  $T_{\#}\mu$  on  $(Y, \mathcal{N})$  defined by

$$T_{\#}\mu(B) = \mu(T^{-1}(B)) \quad \text{for all } B \in \mathcal{N}.$$

- **Absolutely continuous functions.**  $AC(I)$  denotes the set of *absolutely continuous functions* on the interval  $I$ .
- **Characteristic function of a set.** The *characteristic function* of a set  $A$  is defined as

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

- **Dirac mass.** Let  $x \in \mathbb{R}^n$ . The *Dirac mass* in  $x$  is the distribution  $\delta_x$  defined by

$$\langle \delta_x, \varphi \rangle = \varphi(x) \quad \text{for all } \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$$

with  $\mathcal{C}_c^\infty(\mathbb{R}^n)$  the compactly supported smooth functions on  $\mathbb{R}^n$ . Equivalently, it is the measure with  $\delta_x(A) = 1$  if  $x \in A$  and  $\delta_x(A) = 0$  else.

In the upper and lower bound proofs we will furthermore employ the following abbreviations.

- **Cross section.** For every  $0 \leq t \leq 1$ , the set  $\{x_n = t\} = \{x \in \mathbb{R}^n : x_n = t\}$  will be called a *cross-section*. We will also use analogous notations such as  $\{x_n < t\}$  with the obvious meaning.
- **Projection on the hyperplane  $\{x_n = 0\}$ .** By  $x'$  we mean the *projection* of  $x = (x_1, \dots, x_n)$  on the hyperplane  $\{x_n = 0\}$ ,  $x' = (x_1, \dots, x_{n-1}, 0)$ .
- **Tubular neighbourhood.**  $B_s(U)$  is the *tubular neighbourhood* of radius  $s$  of the set  $U \subset \mathbb{R}^n$ ,

$$B_s(U) = \{x \in \mathbb{R}^n : \inf_{y \in U} |x - y| \leq s\}.$$

- $C \equiv C(n)$  identifies  $C$  as a constant depending only on the dimension  $n$ ; if the constant depends on other parameters as well, an analogous notation is used.
- When we write  $A \lesssim B$  or  $A \gtrsim B$ , we mean that there exists a constant  $C \equiv C(n)$  such that  $A \leq CB$  or  $B \leq CA$ , respectively.  $A \sim B$  stands for  $A \lesssim B$  and  $A \gtrsim B$ .
- $\omega_n$  denotes the  $n$ -dimensional measure  $\mathcal{L}^n(B_1(0))$  of the unit ball in  $\mathbb{R}^n$ .

Next we introduce briefly the *Wasserstein distance* and *Kantorovich-Rubinstein duality*. We refer to [Vil03, Chap. 1] for a complete account on Wasserstein distances, spaces and their properties.

**Definition 1.1.1** (Wasserstein distance). Let  $\mu_+, \mu_-$  be finite Borel measures with equal mass  $\|\mu_+\|_{\text{fbm}} = \|\mu_-\|_{\text{fbm}}$ . Given  $p \geq 1$ , the  $p^{\text{th}}$  *Wasserstein distance* between  $\mu_+$  and  $\mu_-$  is

$$W_p(\mu_+, \mu_-) = \left( \inf_{\mu \in \Pi(\mu_+, \mu_-)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^p d\mu(x, y) \right)^{\frac{1}{p}},$$

where  $\Pi(\mu_+, \mu_-) = \{\mu \in \text{fbm}(\mathbb{R}^n \times \mathbb{R}^n) : \mu(A \times \mathbb{R}^n) = \mu_+(A), \mu(\mathbb{R}^n \times B) = \mu_-(B) \text{ for all Borel sets } A, B \subset \mathbb{R}^n\}$ .

We recall the following theorem, which relates the optimal transport problem with its dual formulation (see [Vil03, Thm. 1.14]).

**Theorem 1.1.2** (Kantorovich-Rubinstein duality). *Let  $\mu_+, \mu_- \in \text{fbm}(\mathbb{R}^n)$  with equal mass. Then*

$$W_1(\mu_+, \mu_-) = \sup \left\{ \int_{\mathbb{R}^n} \varphi d(\mu_+ - \mu_-) : \varphi : \mathbb{R}^n \rightarrow \mathbb{R} \text{ Lipschitz with constant } 1 \right\}.$$

Finally, for the reader's convenience we compile here a reference list of the most important symbols with references to the corresponding definitions.

- $I = [0, 1]$ : The unit interval.
- $d_\Sigma$ : Urban planning transport metric (see Formula (2.2.1)).
- $(\Gamma, \mathcal{B}(\Gamma), P_\Gamma)$ : Reference space of all particles (Definition 2.1.1).
- $\chi$ : Irrigation pattern of all particles (Definition 2.1.3).
- $[x]_\chi$ : Solidarity class of  $x$  (Definition 2.1.5).
- $\mu_+^\chi, \mu_-^\chi$ : Irrigating and irrigated measure (Definition 2.1.4).
- $t_p(s)$  and  $\xi^*(p, s)$ : Crossing time and crossing point for particle  $p$  and cross-section  $\{x_n = s\}$  (see Formulae (4.2.2) and (4.2.3)).

## 2 Reminder of transport networks and energy scaling

In this section we introduce different formulations of branched transport and urban planning which prove beneficial in deriving energy scaling laws.

### 2.1 Irrigation patterns

As a preparation we here recall the Lagrangian or pattern-based formulation of particle transport (see [MSM03], [BCM05], [MS13]).

**Definition 2.1.1** (Reference space). Consider a complete separable uncountable metric space  $\Gamma$  endowed with the  $\sigma$ -algebra  $\mathcal{B}(\Gamma)$  of its Borel sets and a positive finite Borel measure  $P_\Gamma$  with no atoms. We refer to  $(\Gamma, \mathcal{B}(\Gamma), P_\Gamma)$  as the *reference space*.

The reference space can be interpreted as the space of all particles that will be transported from a distribution  $\mu_+$  to a distribution  $\mu_-$ .

*Remark 2.1.2.* Any reference space can be shown to be isomorphic to the *standard space*  $([0, 1], \mathcal{B}([0, 1]), m\mathcal{L}^1 \llcorner [0, 1])$  with  $m = P_\Gamma(\Gamma)$  (see [Roy88, Prop. 12 or Thm. 16 in Sec. 5 of Chap. 15] or [Vil09, Chap. 1] for a proof). As a consequence, the following definitions and results are independent of the particular choice of the reference space, and we will just assume it to be the standard space unless stated otherwise.

**Definition 2.1.3** (Irrigation pattern). Let  $I = [0, 1]$  and  $(\Gamma, \mathcal{B}(\Gamma), P_\Gamma)$  be our reference space. An *irrigation pattern* is a measurable function  $\chi : \Gamma \times I \rightarrow \mathbb{R}^n$  such that for almost all  $p \in \Gamma$  we have  $\chi_p \in AC(I)$ .

A pattern  $\tilde{\chi}$  is *equivalent* to  $\chi$  if the images of  $P_\Gamma$  through the maps  $p \mapsto \chi_p, p \mapsto \tilde{\chi}_p$  are the same.

For intuition,  $\chi_p$  can be viewed as the path of particle  $p$ . Its image  $\chi_p(I)$  is called a *fibre* and will frequently be identified with the particle  $p$ .

Here we followed the setting recently introduced in [MS13].

**Definition 2.1.4** (Irrigating and irrigated measure). Let  $\chi$  be an irrigation pattern. Let  $i_0^\chi, i_1^\chi : \Gamma \rightarrow \mathbb{R}^n$  be defined as  $i_0^\chi(p) = \chi(p, 0)$  and  $i_1^\chi(p) = \chi(p, 1)$ . The *irrigating measure* and the *irrigated measure* are defined as the pushforward of  $P_\Gamma$  via  $i_0^\chi$  and  $i_1^\chi$ , respectively:  $\mu_+^\chi = (i_0^\chi)_\# P_\Gamma$ ,  $\mu_-^\chi = (i_1^\chi)_\# P_\Gamma$ .

**Definition 2.1.5** (Solidarity class). For every  $x \in \mathbb{R}^n$ , the set of all particles flowing through  $x$  is denoted

$$[x]_\chi = \{q \in \Gamma : x \in \chi_q(I)\}$$

and its mass by  $m_\chi(x) = P_\Gamma([x]_\chi)$ .

## 2.2 An urban planning model

As our first model of transport networks we consider the so-called urban planning model from [BB05]. Its original motivation is to optimise the public transport network for the daily commute of employees from their homes (described by a spatial distribution  $\mu_0 \in \text{fbm}(\mathbb{R}^n)$ ) to their workplaces (described by  $\mu_1$ ).

The public transport network is represented by a rectifiable set  $\Sigma \subset \mathbb{R}^n$ , and a person travelling along a path  $\theta \subset \mathbb{R}^n$  has expenses

$$a\mathcal{H}^1(\theta \setminus \Sigma) + b\mathcal{H}^1(\theta \cap \Sigma),$$

where  $a$  and  $b < a$  denote the cost per travelling distance outside and on the network, respectively. The minimum cost to get from  $x \in \mathbb{R}^n$  to  $y \in \mathbb{R}^n$  is thus given by

$$d_\Sigma(x, y) = \inf\{a\mathcal{H}^1(\theta \setminus \Sigma) + b\mathcal{H}^1(\theta \cap \Sigma) : \theta \in C_{x,y}\} \quad (2.2.1)$$

for  $C_{x,y} = \{\theta : [0, 1] \rightarrow \mathbb{R}^n : \theta \text{ is Lipschitz with } \theta(0) = x, \theta(1) = y\}$ . The objective functional describing the cost efficiency of the network  $\Sigma$  is now given as

$$\mathcal{E}_W^{\varepsilon,a,b,\mu_0,\mu_1}[\Sigma] = W_{d_\Sigma}(\mu_0, \mu_1) + \varepsilon\mathcal{H}^1(\Sigma),$$

where  $\varepsilon\mathcal{H}^1(\Sigma)$  models the maintenance cost for the network and the Wasserstein-type distance

$$W_{d_\Sigma}(\mu_0, \mu_1) = \inf_{\mu \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d_\Sigma(x, y) d\mu(x, y)$$

with  $\Pi(\mu_0, \mu_1) = \{\mu \in \text{fbm}(\mathbb{R}^n \times \mathbb{R}^n) : \pi_{1\#}\mu = \mu_0, \pi_{2\#}\mu = \mu_1\}$  represents the total travel expenses of the population travelling from their homes  $\mu_0$  to their workplaces  $\mu_1$ . Due to  $\mathcal{E}_W^{\varepsilon,a,b,\mu_0,\mu_1} = b\mathcal{E}_W^{\frac{\varepsilon}{b},\frac{a}{b},1,\mu_0,\mu_1}$  it actually suffices to study the case  $b = 1$ .

In [BW15] it has been shown that the urban planning model can equivalently be reformulated in terms of irrigation patterns  $\chi$  as detailed below. The optimal transport network  $\Sigma^*$  can then be extracted from the optimal irrigation pattern  $\chi^*$  via  $\Sigma^* = \{x \in \mathbb{R}^n : m_\chi^*(x) > \frac{\varepsilon}{a-1}\}$ . The pattern formulation is advantageous to prove energy scaling laws, which is why we introduce it here.

**Definition 2.2.1** (Urban planning problem, pattern formulation). Let  $(\Gamma, \mathcal{B}(\Gamma), P_\Gamma)$  be the reference space and  $\chi : \Gamma \times [0, 1] \rightarrow \mathbb{R}^n$  be an irrigation pattern. For  $\varepsilon > 0$  and  $a > 1$ , consider the cost density

$$r_{\varepsilon,a}^\chi(x) = \begin{cases} \min\left\{1 + \frac{\varepsilon}{m_\chi(x)}, a\right\} & \text{if } m_\chi(x) > 0, \\ a & \text{if } m_\chi(x) = 0. \end{cases}$$

The *urban planning cost functional*  $E_P^{\varepsilon,a}$  is given by

$$E_P^{\varepsilon,a}(\chi) = \int_{\Gamma \times I} r_{\varepsilon,a}^\chi(\chi_p(t)) |\dot{\chi}_p(t)| dP_\Gamma(p) dt.$$

For  $\mu_0, \mu_1 \in \text{fbm}(\mathbb{R}^n)$ , the *urban planning problem* is

$$\min \mathcal{E}^{\varepsilon,a,\mu_0,\mu_1}[\chi],$$

where

$$\mathcal{E}^{\varepsilon,a,\mu_0,\mu_1}[\chi] = \begin{cases} E_P^{\varepsilon,a}(\chi) & \text{if } \mu_+^\chi = \mu_0 \text{ and } \mu_-^\chi = \mu_1, \\ \infty & \text{else.} \end{cases}$$

Existence of minimising patterns is shown in [BW15] as well as the equivalence of the optimisation problems  $\min_\chi \mathcal{E}^{\varepsilon,a,\mu_0,\mu_1}[\chi] = \min_\Sigma \mathcal{E}_W^{\varepsilon,a,1,\mu_0,\mu_1}[\Sigma]$ . Note that  $\varepsilon$  governs the network maintenance costs. Thus, a large  $\varepsilon$  will lead to coarse networks, while a small  $\varepsilon$  allows the network to become very complex. Our energy scaling law will be concerned with the latter case.

Finally, we will later need the following reparameterisation result from [BW15].

**Proposition 2.2.2** (Constant speed reparameterisation of patterns). *Irrigation patterns of finite cost can be reparameterised such that  $\chi_p : I \rightarrow \mathbb{R}^n$  is Lipschitz and  $|\dot{\chi}_p|$  is constant for almost all  $p \in \Gamma$  without changing the cost  $E_P^{\varepsilon,a}$ .*

## 2.3 Branched transport

The second model considered is the *branched transportation problem*. In the literature, the branched transport functional is commonly parameterised by a parameter  $\alpha \in [0, 1)$ , which describes the economies of scale for mass transport: The cost for transporting mass  $m$  by a unit distance is  $m^\alpha$ , so the cost per single mass particle gets smaller the more particles are transported together. This creates a preference for branching networks in which mass is first lumped together and then transported in bulk. The closer  $\alpha$  to one, the less pronounced is this preference and the more complex the transportation networks can become. Since we will study the asymptotic behaviour of finer and finer networks, we will in our description replace  $\alpha$  by  $\varepsilon = 1 - \alpha$  and consider the case of  $\varepsilon$  small. The following definition of the functional  $M_P^\varepsilon$  follows [BCM05].

**Definition 2.3.1** (Branched transport problem). For  $0 < \varepsilon \leq 1$ , the reference space  $(\Gamma, \mathcal{B}(\Gamma), P_\Gamma)$ , and an irrigation pattern  $\chi : \Gamma \times [0, 1] \rightarrow \mathbb{R}^n$  we consider the *cost density*

$$s_\varepsilon^\chi(x) = \begin{cases} m_\chi(x)^{-\varepsilon} & \text{if } m_\chi(x) > 0, \\ \infty & \text{if } m_\chi(x) = 0. \end{cases}$$

The *branched transport cost functional* associated with irrigation pattern  $\chi$  is

$$M_P^\varepsilon(\chi) = \int_{\Gamma \times I} s_\varepsilon^\chi(\chi_p(t)) |\dot{\chi}_p(t)| dP_\Gamma(p) dt.$$

For  $\mu_0, \mu_1 \in \text{fbm}(\mathbb{R}^n)$ , the *branched transport problem* is

$$\min \mathcal{M}^{\varepsilon, \mu_0, \mu_1}[\chi],$$

where

$$\mathcal{M}^{\varepsilon, \mu_0, \mu_1}[\chi] = \begin{cases} M_P^\varepsilon(\chi) & \text{if } \mu_+^\chi = \mu_0 \text{ and } \mu_-^\chi = \mu_1, \\ \infty & \text{else.} \end{cases}$$

Given  $\mu_0, \mu_1 \in \text{fbm}(\mathbb{R}^n)$  with compact support, an optimal irrigation pattern exists [MS09].

Finally, we will need the fact from [BW15] that patterns can be reparameterised fibrewise without changing the cost.

**Proposition 2.3.2** (Constant speed reparameterisation of patterns). *Irrigation patterns of finite cost can be reparameterised such that  $\chi_p : I \rightarrow \mathbb{R}^n$  is Lipschitz and  $|\dot{\chi}_p|$  is constant for almost all  $p \in \Gamma$  without changing the cost  $M_P^\varepsilon$ .*

## 2.4 The optimal energy scaling

Recall from Section 1 that for  $A \subset \mathbb{R}^{n-1}$  a hypersquare we will use

$$\mu_0 = m\mathcal{H}^{n-1} \llcorner (A \times \{0\}), \quad \mu_1 = m\mathcal{H}^{n-1} \llcorner (A \times \{L\})$$

Therefore we set  $\mathcal{E}^{\varepsilon, a, A, m, L} = \mathcal{E}^{\varepsilon, a, \mu_0, \mu_1}$  and  $\mathcal{M}^{\varepsilon, A, m, L} = \mathcal{M}^{\varepsilon, \mu_0, \mu_1}$ . By rescaling the geometry and the mass flux one easily arrives at the following non-dimensionalisations,

$$\begin{aligned} \mathcal{E}^{\varepsilon, a, A, m, L}[\chi] &= L^n m \mathcal{E}^{\frac{\varepsilon}{L^{n-1}m}, a, \frac{1}{L}A, 1, 1}[\frac{1}{L}\chi], \\ \mathcal{M}^{\varepsilon, A, m, L}[\chi] &= (L^n m)^{1-\varepsilon} L \mathcal{M}^{\varepsilon, \frac{1}{L}A, 1, 1}[\frac{1}{L}\chi], \end{aligned}$$

so that without loss of generality we may restrict ourselves to the study of

$$\begin{aligned} \mu_0 &= \mathcal{H}^{n-1} \llcorner (A \times \{0\}), \quad \mu_1 = \mathcal{H}^{n-1} \llcorner (A \times \{1\}), \\ \mathcal{E}^{\varepsilon, a, A} &:= \mathcal{E}^{\varepsilon, a, A, 1, 1} \quad \text{and} \quad \mathcal{M}^{\varepsilon, A} := \mathcal{M}^{\varepsilon, A, 1, 1}. \end{aligned}$$

Furthermore, let us denote the minimum energy for  $\varepsilon = 0$  by

$$\mathcal{E}^{*, a, A} = \inf_{\chi} \mathcal{E}^{0, a, A}[\chi], \quad \mathcal{M}^{*, A} = \inf_{\chi} \mathcal{M}^{0, A}[\chi].$$

We have

$$\mathcal{E}^{*,a,A} = \mathcal{M}^{*,A} = W_1(\mu_0, \mu_1) = \mathcal{H}^{n-1}(A)$$

for  $W_1(\mu_0, \mu_1) = \inf_{\mu \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\mu(x, y)$  the Wasserstein distance. Indeed, our constructions in Section 3 will imply  $\mathcal{M}^{*,A}, \mathcal{E}^{*,a,A} \leq W_1(\mu_0, \mu_1)$ , while Sections 4.2 and 4.3 begin with calculations showing  $\mathcal{M}^{*,A}, \mathcal{E}^{*,a,A} \geq W_1(\mu_0, \mu_1)$ . For now, the reader may simply imagine the infimum value being approached by irrigation patterns  $\chi : \Gamma \times I \rightarrow \mathbb{R}^n$  with  $\Gamma = A$  approximating  $\chi_p(t) = (p_1, \dots, p_{n-1}, t)^T$ , transport through a network with infinitely many vertical pipes of length 1.

For  $\varepsilon > 0$  the minimum achievable energy deviates from  $\mathcal{M}^{*,A}$  and  $\mathcal{E}^{*,a,A}$ , respectively. We will call this deviation the excess energy and abbreviate it by

$$\Delta \mathcal{E}^{\varepsilon,a,A} = \min_{\chi} \mathcal{E}^{\varepsilon,a,A}[\chi] - \mathcal{E}^{*,a,A} \quad \text{and} \quad \Delta \mathcal{M}^{\varepsilon,A} = \min_{\chi} \mathcal{M}^{\varepsilon,A}[\chi] - \mathcal{M}^{*,A}.$$

The main results of this article are the following two theorems, which show how the deviation scales in  $\varepsilon$ . The first theorem treats the urban planning model.

**Theorem 2.4.1** (Energy scaling for the Urban Planning). *There are constants  $C_1, C_2 > 0$  independent of  $\varepsilon, a, A$  such that for  $\varepsilon < \min\{1, \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}\}$  and  $a > 1$  we have*

$$C_1 \mathcal{H}^{n-1}(A) \min\{a - 1, f(\varepsilon, a)\} \leq \Delta \mathcal{E}^{\varepsilon,a,A} \leq C_2 \mathcal{H}^{n-1}(A) \min\{a - 1, f(\varepsilon, a)\},$$

where the function  $f$  is given by

$$\begin{aligned} f(\varepsilon, a) &= \varepsilon^{\frac{2}{3}} & \text{if } n = 2, \\ f(\varepsilon, a) &= (\sqrt{a} + |\log \frac{a-1}{\sqrt{\varepsilon}}|) \sqrt{\varepsilon} & \text{if } n = 3, \\ f(\varepsilon, a) &= \sqrt{a}(\sqrt{a-1})^{\frac{n-3}{n-1}} \varepsilon^{\frac{1}{n-1}} & \text{if } n > 3. \end{aligned}$$

The proof is given in Sections 3.3 and 4.2. Note that the dependence on the parameter  $a$  is also identified, which plays a dominant role if either  $a$  is huge or close to 1. The three different scaling laws for different dimensions indicate three different regimes, as will become clear in Sections 3.3 and 4.2:

- In 2D, the dominant contribution to the excess energy stems from the interior region between and bounded away from  $\mu_0$  and  $\mu_1$ .
- In 3D, the dominant energy contribution also occurs in between the two measures, but this time it is more evenly distributed, up to the boundary.
- In higher dimensions, the energy scaling is dominated by energy contributions concentrated at  $\mu_0$  and  $\mu_1$ .

The second result is concerned with the branched transport model.

**Theorem 2.4.2** (Energy scaling for the Branched Transport). *There are constants  $C_1, C_2, C_3, C_4 > 0$  independent of  $\varepsilon$  and  $A$  such that for  $\varepsilon < \min\{\mathcal{H}^{n-1}(A)^{\frac{2}{n-1}}, \mathcal{H}^{n-1}(A)^{-C_3}, C_4\}$  we have*

$$C_1 \mathcal{H}^{n-1}(A) \varepsilon |\log \varepsilon| \leq \Delta \mathcal{M}^{\varepsilon,A} \leq C_2 \mathcal{H}^{n-1}(A) \varepsilon |\log \varepsilon|.$$

The proof is given in Sections 3.4 and 4.3 and reveals that the total energy distribution follows a similar pattern as in 3D urban planning.

All above energy scalings will be obtained following variations of a general scheme originally introduced by Kohn and Müller in 1992 [KM92] to model martensite–austenite interfaces.

- The upper bound is based on a construction with unit cells arranged in hierarchical layers. The unit cells of the layer closest to the domain boundary are the smallest, and from layer to layer the unit cell width doubles, thereby leading to a coarsening towards the domain centre. Each single unit cell contains a branching structure compatible with the width doubling between two layers (in Figure 1, bottom left, for instance, a unit cell is a ‘V’-shaped structure). The free design parameters are the unit cell widths on the coarsest and the finest level as well as the aspect ratio of the unit cells on each level. Now the excess energy due to  $\varepsilon > 0$  can be computed for each unit cell as a function of the design parameters, and their sum yields the total excess energy. Minimisation for the design parameters yields the optimal construction and energy scaling. The detailed procedure is provided in Sections 3.3 and 3.4.



- The technique for the lower bound exploits the fact that the transport costs between any two measures for  $\varepsilon > 0$  can be bounded below by the transport costs for  $\varepsilon = 0$ , which can be calculated by solving a (much simpler) convex problem. In essence, one first characterises a generic cross-section in terms of how many interfaces or pipes it intersects at most. The number depends on  $\varepsilon$  and usually is a simple consequence of the excess energy being bounded above. Then, the energy can be bounded below by the costs one would have for  $\varepsilon = 0$ , knowing the additional information about the generic cross-section. Those costs can be obtained by solving a *convex* optimisation problem due to  $\varepsilon = 0$ . This idea in its plainest form is exemplified in Section 4.2 for dimension  $n = 2$ , and variants of it are applied in Section 4.2 for  $n \geq 3$  and in Section 4.3.

### 3 Upper bound via branching construction

In this section we will provide constructions of transport networks with the required energy scaling. Those constructions will be composed of elementary units, all described as an irrigation pattern, for which reason we first define several operations how multiple irrigation patterns can be combined to a single one in Section 3.1. Afterwards we define and analyse the single elementary units in Section 3.2 as well as the optimal constructions for urban planning (Section 3.3) and branched transport (Section 3.4).

#### 3.1 Operations with patterns

Our constructions will require a combination of irrigation patterns in parallel as well as in series. Those operations are defined in Definition 3.1.1 and 3.1.2, respectively. Definition 3.1.1 was introduced in [BCM08] and has also been given in an equivalent form in [BCM09, Lemma 5.5]. Here the particles of all constituting subpatterns are kept independent.

**Definition 3.1.1** (Union of patterns). Let  $\chi_i : \Gamma_i \times I \rightarrow \mathbb{R}^n$ ,  $i \in N \subset \mathbb{N}$ , be a sequence of irrigation patterns with reference spaces  $(\Gamma_i, \mathcal{B}(\Gamma_i), P_{\Gamma_i})$ . Suppose that  $\sum_i P_{\Gamma_i}(\Gamma_i) < +\infty$ . Let  $\Gamma$  be the disjoint union of the  $\Gamma_i$ ,

$$\Gamma = \coprod_{i \in N} \Gamma_i,$$

endowed with the product topology and the corresponding Borel  $\sigma$ -algebra. Furthermore consider the measure  $P_\Gamma$  defined by

$$P_\Gamma(A) = \sum_{i \in N} P_{\Gamma_i}(A \cap \Gamma_i).$$

The irrigation pattern  $\chi : \Gamma \times I \rightarrow \mathbb{R}^n$ , defined by

$$\chi(p, t) = \chi_i(p, t), \text{ if } p \in \Gamma_i,$$

is called the *union* of the patterns  $\chi_i$  and will be denoted by  $\chi = \coprod_i \chi_i$ .

Next we join two patterns  $\chi_1$  and  $\chi_2$  in series. Here, every particle  $p$  first travels along a path defined by  $\chi_1$  and then continues along a path defined by  $\chi_2$ .

**Definition 3.1.2** (Series of patterns). Let  $\chi_i : \Gamma_i \times I \rightarrow \mathbb{R}^n$  for  $i = 1, 2$  be two irrigation patterns and  $T : \Gamma_1 \rightarrow \Gamma_2$  an isomorphism of measure spaces such that

$$\chi_1(p, 1) = \chi_2(T(p), 0) \quad \text{for } P_\Gamma\text{-almost all } p \in \Gamma_1.$$

Now define the irrigation pattern  $\tilde{\chi} : \Gamma_1 \times I \rightarrow \mathbb{R}^n$  by

$$\tilde{\chi}(p, t) = \begin{cases} \chi_1(p, 2t) & t \in [0, \frac{1}{2}], \\ \chi_2(T(p), 2t - 1) & t \in [\frac{1}{2}, 1]. \end{cases}$$

Furthermore, for  $p \in \Gamma_1$  let  $r_p : I \rightarrow I$  be a reparameterisation such that  $\chi(p, t) = \tilde{\chi}(p, r_p(t))$  has constant speed  $|\dot{\chi}_p|$  for almost all  $p \in \Gamma_1$ . The irrigation pattern  $\chi$  is called the *T-relative series* of patterns  $\chi_1$  and  $\chi_2$  and is denoted by  $\chi_1 \circ_T \chi_2$ .

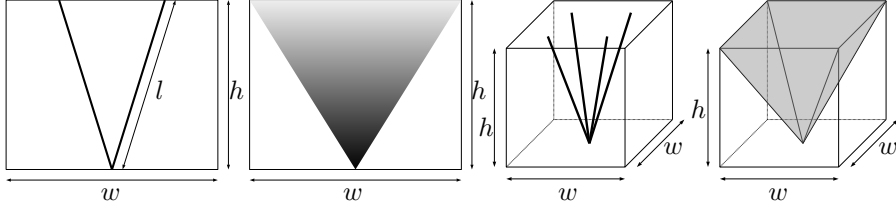


Figure 2: Elementary and Wasserstein cells in 2D and 3D.

*Remark 3.1.3.* A similar construction was given in [BCM09, Lemma 5.5]. Ours is a little more technical since we track the dependence on  $T$ , while in [BCM09] the authors just require  $\mu_-^{\chi_1} = \mu_+^{\chi_2}$  and from this anticipate the existence of some measure-preserving map  $T : \Gamma_1 \rightarrow \Gamma_2$  with  $\chi_1(p, 1) = \chi_2(T(p), 0)$ , which is then used in the construction. However, such a map  $T$  may not exist, as shown by the example  $\Gamma_1 = \Gamma_2 = [0, 1]$  with  $P_\Gamma = \mathcal{L}^1 \llcorner [0, 1]$ ,  $\chi_1(p, 1) = 2p \bmod 1$ ,  $\chi_2(p, 0) = p$ . The result in [BCM09] certainly stays valid, but its proof will require to relax the transport map  $T$  between particles to a transport plan. To circumvent those difficulties we shall simply always specify the map  $T$  explicitly.

*Remark 3.1.4.* It is straightforward to verify the following elementary properties, in which  $\mathcal{J}$  either stands for the urban planning energy  $E_P^{\varepsilon, a}$  or the branched transport energy  $M_P^\varepsilon$ ,

- $\mathcal{J}(\chi_1 \amalg \chi_2) \leq \mathcal{J}(\chi_1) + \mathcal{J}(\chi_2)$ ,
- $\mathcal{J}(\chi_1 \circ_T \chi_2) \leq \mathcal{J}(\chi_1) + \mathcal{J}(\chi_2)$ ,
- $\mu_+^{\chi_1 \amalg \chi_2} = \mu_+^{\chi_1} + \mu_+^{\chi_2}$ ,  $\mu_-^{\chi_1 \amalg \chi_2} = \mu_-^{\chi_1} + \mu_-^{\chi_2}$ ,
- $\mu_+^{\chi_1 \circ_T \chi_2} = \mu_+^{\chi_1}$ ,  $\mu_-^{\chi_1 \circ_T \chi_2} = \mu_-^{\chi_2}$ ,

where the second inequality uses Propositions 2.2.2 and 2.3.2 and where both inequalities become equalities if the patterns do not overlap.

### 3.2 Elementary cells

We introduce here two elementary irrigation patterns which we will use to build more complicated ones. They are illustrated in Figure 2.

**Definition 3.2.1** (Elementary cell in  $nD$ ). Let  $\Gamma_{-1} = [-1, 0]$ ,  $\Gamma_{+1} = (0, 1]$ , and define for a multiindex  $j \in \{-1, 1\}^{n-1}$  the unit square

$$\Gamma_j = \bigtimes_{i=1}^{n-1} \Gamma_{j_i}$$

in the orthant defined by  $j$ . Its midpoint shall be denoted by  $m_j = \frac{1}{2}j \in \mathbb{R}^{n-1}$ .

The *elementary cell* with base point  $x = (x', x_n)$ , width  $w$ , height  $h$ , and flux  $f$  is defined as the triple  $(\Gamma_f, P_{\Gamma_f}, \chi_{x,w,h,f}^E)$  for the reference space  $\Gamma_f = [-1, 1]^{n-1} = \bigcup_{j \in \{-1, 1\}^{n-1}} \Gamma_j$ , the reference measure  $P_{\Gamma_f} = f \mathcal{L}^{n-1} \llcorner \Gamma_f$ , and the irrigation pattern  $\chi_{x,w,h,f}^E : \Gamma_f \times I \rightarrow \mathbb{R}^n$ ,

$$\chi_{x,w,h,f}^E(p, t) = (x' + t \frac{w}{2} m_j, x_n + th) \quad \text{for } p \in \Gamma_j.$$

The *initial* and *final measure* of the elementary cell are

$$\mu_+^{\chi_{x,w,h,f}^E} = 2^{n-1} f \delta_x, \quad \mu_-^{\chi_{x,w,h,f}^E} = f \sum_{j \in \{-1, 1\}^{n-1}} \delta_{(x' + \frac{w}{2} m_j, x_n + h)}.$$

Note that the image of  $\chi_{x,w,h,f}^E$  consists of  $2^{n-1}$  branches with length

$$l = \sqrt{(n-1) \left(\frac{w}{4}\right)^2 + h^2} = \sqrt{\frac{n-1}{16} w^2 + h^2} \quad (3.2.1)$$

and is fully contained in

$$E = x + \left( \left[ -\frac{w}{2}, \frac{w}{2} \right]^{n-1} \times [0, h] \right).$$

The initial and final measure of the elementary cell are concentrated on the bottom and top boundary of  $E$ , respectively.

**Lemma 3.2.2** (Elementary cell cost). *The urban planning and branched transport energy of an elementary cell satisfy*

$$\begin{aligned} E_P^{\varepsilon, a}(\chi_{x, w, h, f}^E) &= 2^{n-1} \min\{af, f + \varepsilon\}l \\ M_P^{\varepsilon}(\chi_{x, w, h, f}^E) &= 2^{n-1} f^{1-\varepsilon}l \end{aligned}$$

for  $l$  from (3.2.1).

*Proof.* This is a straightforward calculation based on inserting  $\chi_{x, w, h, f}^E$  and  $P_{\Gamma_f}$  into Definition 2.2.1 and Definition 2.3.1, using that the mass through each point  $\tilde{x}$  of a branch is given by  $m_{\chi_{x, w, h, f}^E}(\tilde{x}) = f$ .  $\square$

**Definition 3.2.3** (Wasserstein cell in  $nD$ ). The *Wasserstein cell* with base point  $x = (x', x_n)$ , width  $w$ , height  $h$ , and flux  $f$  is defined as the triple  $(\Gamma_f, P_{\Gamma_f}, \chi_{x, w, h, f}^W)$  for  $\Gamma_f, P_{\Gamma_f}$  as in the previous definition and the irrigation pattern  $\chi_{x, w, h, f}^W : \Gamma_f \times I \rightarrow \mathbb{R}^n$ ,

$$\chi_{x, w, h, f}^W(p, t) = \left( x' + t \frac{w}{2} p, x_n + th \right).$$

The *initial* and *final measure* of the Wasserstein cell are

$$\mu_+^{\chi_{x, w, h, f}^W} = 2^{n-1} f \delta_x, \quad \mu_-^{\chi_{x, w, h, f}^W} = f \left( \frac{w}{2} \right)^{1-n} \mathcal{H}^{n-1} \llcorner \left( x + \left( \left[ -\frac{w}{2}, \frac{w}{2} \right]^{n-1} \times \{h\} \right) \right).$$

Just like the elementary cell, the Wasserstein cell is fully contained in  $E$ .

**Lemma 3.2.4** (Wasserstein cell urban planning cost). *The urban planning energy of a Wasserstein cell satisfies*

$$E_P^{\varepsilon, a}(\chi_{x, w, h, f}^W) \leq \begin{cases} \min\{af, f + \varepsilon\}w & \text{if } n = 2 \text{ and } h = 0, \\ 2^{n-1} af \sqrt{\frac{(n-1)}{4} w^2 + h^2} & \text{else.} \end{cases}$$

*Proof.* If  $n \neq 2$  or  $h \neq 0$  we have  $m_{\chi_{x, w, h, f}^W}(\tilde{x}) = 0$  for all  $\tilde{x} \neq x$ . Using this in Definition 2.2.1 we obtain

$$\begin{aligned} E_P^{\varepsilon, a}(\chi_{x, w, h, f}^W) &= \int_{\Gamma_f} \int_0^1 r_{\varepsilon, a}^{\chi}(\chi_{x, w, h, f}^W(p, t)) |\dot{\chi}_{x, w, h, f}^W(p, t)| dt dP_{\Gamma_f}(p) \\ &= a \int_{\Gamma_f} \sqrt{\frac{w^2}{4} |p|^2 + h^2} dP_{\Gamma_f}(p) \leq 2^{n-1} af \sqrt{\frac{(n-1)}{4} w^2 + h^2}. \end{aligned}$$

The case  $n = 2, h = 0$  follows analogously, using  $m_{\chi_{x, w, h, f}^W}(\tilde{x}) = f(1 - \frac{|\tilde{x}_1 - x_1|}{w/2})$  for any  $\tilde{x} \in (x_1 - \frac{w}{2}, x_1 + \frac{w}{2}) \times \{x_2\}$ .  $\square$

### 3.3 Upper bound for optimal urban planning energy $\mathcal{E}^{\varepsilon, a, A}$

To derive an upper bound we have to construct an irrigation pattern with the desired energy. It turns out that branched transport in dimension  $n = 2, 3$ , and higher requires slightly different constructions, however, the basic approach is the same for all of them. In particular, we will employ a pattern as illustrated in Figures 3 and 4. It is symmetric about  $x_n = \frac{1}{2}$  so that it suffices to consider the upper half. It consists of  $K$  layers, where the  $k^{\text{th}}$  layer just represents an array of  $N_k^{n-1}$  identical elementary cells. At the top boundary an additional layer of Wasserstein cells is added.

In more detail, without loss of generality we assume the hypersquare  $A$ , whose sidelength is  $\ell = \sqrt[n-1]{\mathcal{H}^{n-1}(A)}$ , to be given by

$$A = [0, \ell]^{n-1}.$$

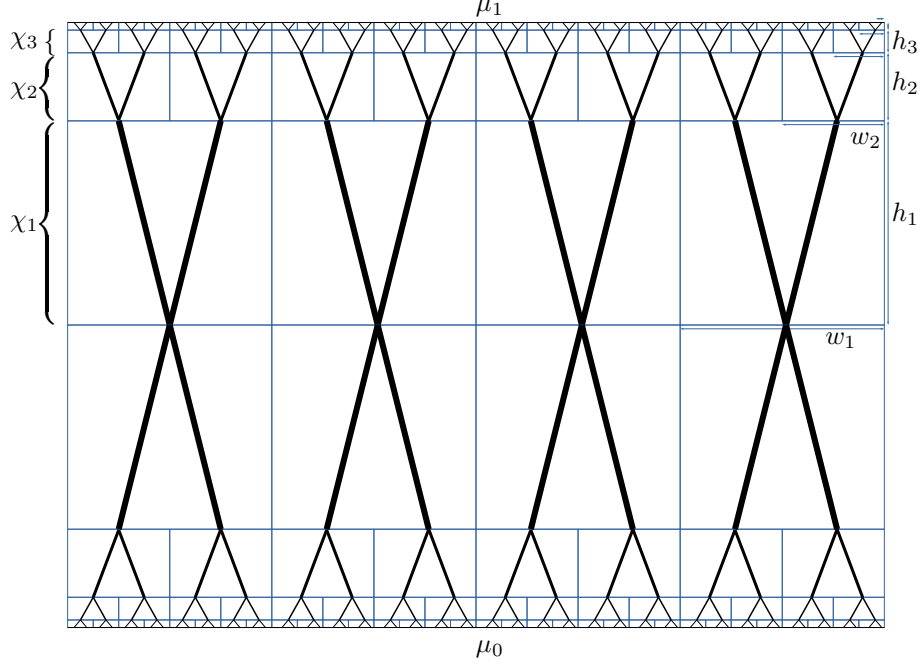


Figure 3: Illustrative sketch of the construction ansatz in two dimensions. The thin lines indicate the boundaries of the elementary cells.

Let  $x_{k,i}$ ,  $w_k$ ,  $h_k$ , and  $f_k$  denote the base point, width, height, and flux associated with the  $i^{\text{th}}$  elementary cell in the  $k^{\text{th}}$  layer, where  $i \in \{1, \dots, N_k\}^{n-1}$  is a multiindex. We choose

$$w_k = 2^{1-k} w_1, \quad N_k = \frac{\ell}{w_k}, \quad f_k = \left(\frac{w_k}{2}\right)^{n-1},$$

$$h_k = c w_k^\alpha \varepsilon^\beta a^\gamma (a-1)^\delta, \quad x_{k,i} = \left( (i_1 - \frac{1}{2}) w_k, \dots, (i_{n-1} - \frac{1}{2}) w_k, \frac{1}{2} + \sum_{j=1}^{k-1} h_j \right) \quad (3.3.1)$$

for the coarsest cell width  $w_1$  and some constants  $c, \alpha, \beta, \gamma, \delta \in \mathbb{R}$  to be determined. Now we define the irrigation pattern describing the  $k^{\text{th}}$  layer as

$$\chi_k = \coprod_{i \in \{1, \dots, N_k\}^{n-1}} \chi_{x_{k,i}, w_k, h_k, f_k}^E$$

and the top layer of Wasserstein cells as

$$\chi_{K+1} = \coprod_{i \in \{1, \dots, N_K\}^{n-1}} \chi_{x_{K+1,i}, w_{K+1}, H, f_{K+1}}^W$$

for some  $K \in \mathbb{N}$ ,  $H \in \mathbb{R}$ . Note that for  $k = 1, \dots, K+1$ , the reference space of  $\chi_k$  is the disjoint union  $\Gamma_k = \coprod_{i \in \{1, \dots, N_k\}^{n-1}} \Gamma_{f_k}^i$ , where  $\Gamma_{f_k}^i$  just denotes a copy of  $\Gamma_{f_k}$ . Now it is straightforward to check that the function

$$T_k : \Gamma_k \rightarrow \mathbb{R}^{n-1}, \quad \Gamma_{f_k}^i \ni p \mapsto x_{k,i} + \frac{w}{2} p$$

represents a measure-preserving map from the original reference space  $\Gamma_k$  onto  $(A, \mathcal{B}(A), \mathcal{L}^{n-1} \llcorner A)$ . Hence, we can identify particles in  $\Gamma_k$  with particles in  $A$  via the map  $T_k$  and thus may take  $(A, \mathcal{B}(A), \mathcal{L}^{n-1} \llcorner A)$  as the reference space of  $\chi_k$ . This simplifies the composition of those patterns, as now we have

$$\chi_k(p, 1) = \chi_{k+1}(p, 0) \quad \text{for almost all } p \in A$$

so that we may define the series of all layers in the upper half as

$$\chi_u = ((\dots ((\chi_1 \circ_{\text{Id}} \chi_2) \circ_{\text{Id}} \chi_3) \circ_{\text{Id}} \dots) \circ_{\text{Id}} \chi_{K+1}),$$

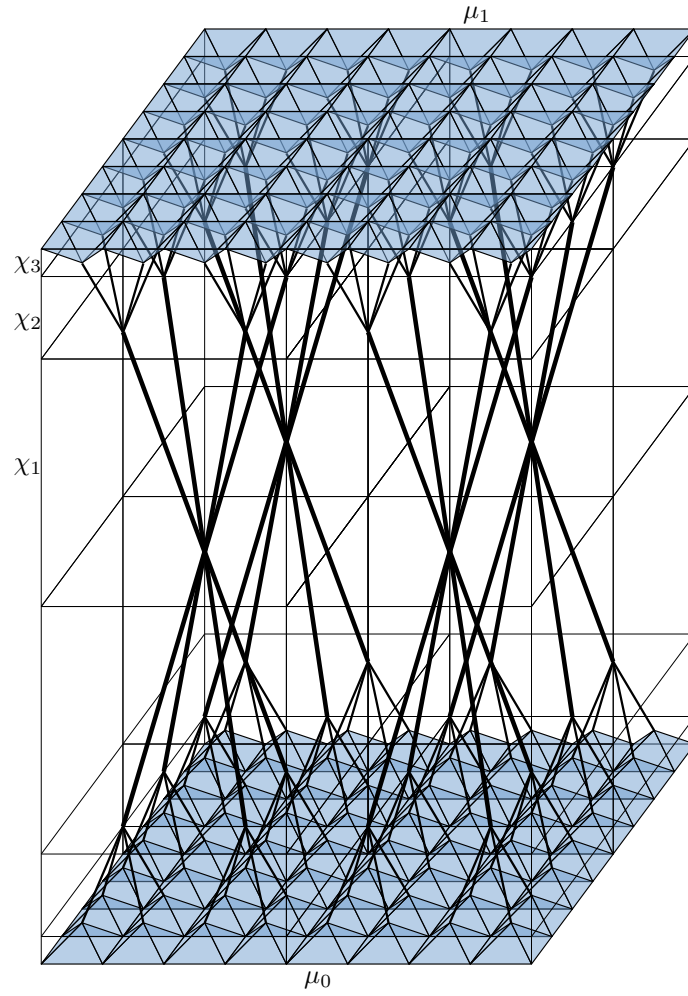


Figure 4: Illustrative sketch of the construction ansatz in three dimensions. The thin lines indicate the boundaries of the elementary cells, the shaded regions represent Wasserstein cells.

using Definition 3.1.2 with the identity map  $\text{Id} : A \rightarrow A$ . The irrigation pattern  $\chi_l$  of the lower half is constructed in the analogous way, and the full irrigation pattern is defined as

$$\chi = \chi_l \circ_{\text{Id}} \chi_u.$$

It turns out that the constraint of the construction having total height 1 already fixes the width of the coarsest elementary cells. Indeed, since the heights of all layers have to add up to 1 we require

$$1 = 2 \left( \sum_{k=1}^K h_k + H \right) = 2c(2w_1)^\alpha \varepsilon^\beta a^\gamma (a-1)^\delta \sum_{k=1}^K 2^{-k\alpha} + 2H = 2cw_1^\alpha \varepsilon^\beta a^\gamma (a-1)^\delta \frac{1-2^{-K\alpha}}{1-2^{-\alpha}} + 2H. \quad (3.3.2)$$

Assuming  $1-2H$  to still be of order 1 it transpires that we should roughly have  $w_1 \sim \varepsilon^{-\frac{\beta}{\alpha}} a^{-\frac{\gamma}{\alpha}} (a-1)^{-\frac{\delta}{\alpha}}$ . Based on these heuristics we simply choose

$$w_1 = \tilde{c} \varepsilon^{-\frac{\beta}{\alpha}} a^{-\frac{\gamma}{\alpha}} (a-1)^{-\frac{\delta}{\alpha}} \quad \text{for } \tilde{c} \in (\tfrac{1}{2}, 1], \quad (3.3.3)$$

where  $\tilde{c}$  accounts for the fact that the total total width  $\ell$  of the geometry has to be an integer multiple of  $w_1$ , that is,  $\tilde{c}$  is determined such that this holds true. Of course, this requires in particular  $w_1 \leq \ell$ , which is why for all our constructions we will have the constraint

$$\varepsilon \leq a^{-\frac{\gamma}{\beta}} (a-1)^{-\frac{\delta}{\beta}} \mathcal{H}^{n-1}(A)^{-\frac{\alpha}{\beta(n-1)}}. \quad (3.3.4)$$

The constant  $c$  is now obtained from (3.3.2) as

$$c = \frac{1-2H}{2\tilde{c}^\alpha} \frac{1-2^{-\alpha}}{1-2^{-K\alpha}} \in [\frac{1-2^{-\alpha}}{4}, 2^{\alpha-1}]. \quad (3.3.5)$$

Here the possible range of  $c$  is based on the conditions

$$K \geq 1, \quad H \leq \tfrac{1}{4}, \quad \alpha > 0, \quad (3.3.6)$$

which will be ensured in all our constructions.

We next have to estimate the energy of  $\chi$ . By the symmetry of the construction and Remark 3.1.4 we have

$$\mathcal{E}^{\varepsilon,a,A}[\chi] = 2E_{\text{P}}^{\varepsilon,a}[\chi_u] = 2 \left[ \sum_{k=1}^K \sum_{i \in \{1, \dots, N_k\}^{n-1}} E_{\text{P}}^{\varepsilon,a}[\chi_{x_k,i,w_k,h_k,f_k}^{\text{E}}] + \sum_{i \in \{1, \dots, N_{K+1}\}^{n-1}} E_{\text{P}}^{\varepsilon,a}[\chi_{x_{K+1},i,w_{K+1},H,f_{K+1}}^{\text{W}}] \right].$$

On the other hand, using  $2 \left( \sum_{k=1}^K h_k + H \right) = 1$  and  $N_k = \frac{\ell}{w_k}$ , the energy value  $\mathcal{E}^{*,a,A}$  can be written as

$$\mathcal{E}^{*,a,A} = W_1(\mu_0, \mu_1) = \mathcal{H}^{n-1}(A) = 2 \left( \sum_{k=1}^K N_k^{n-1} w_k^{n-1} h_k + N_{K+1}^{n-1} w_{K+1}^{n-1} H \right).$$

Thus we obtain

$$\begin{aligned} \Delta \mathcal{E}^{\varepsilon,a,A} &\leq \mathcal{E}^{\varepsilon,a,A}[\chi] - \mathcal{E}^{*,a,A} \\ &\leq 2 \left[ \sum_{k=1}^K \sum_{i \in \{1, \dots, N_k\}^{n-1}} \left( E_{\text{P}}^{\varepsilon,a}[\chi_{x_k,i,w_k,h_k,f_k}^{\text{E}}] - w_k^{n-1} h_k \right) \right. \\ &\quad \left. + \sum_{i \in \{1, \dots, N_{K+1}\}^{n-1}} \left( E_{\text{P}}^{\varepsilon,a}[\chi_{x_{K+1},i,w_{K+1},H,f_{K+1}}^{\text{W}}] - w_{K+1}^{n-1} H \right) \right], \end{aligned} \quad (3.3.7)$$

in which now Lemmas 3.2.2 and 3.2.4 can be used.

### 3.3.1 Upper bound in 2D

The first results, Theorem 3.3.1 and Corollary 3.3.2, are concerned with the two-dimensional situation.

**Theorem 3.3.1** (Partial upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in 2D). *Using the notation from Section 2.2, there exists  $C > 0$  independent of  $\varepsilon, a, A$  such that for all  $\varepsilon < \min\{1, \mathcal{H}^1(A)^3\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^1(A)C\varepsilon^{\frac{2}{3}}.$$

Before proving Theorem 3.3.1 let us note that it directly implies the following corollary.

**Corollary 3.3.2** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in 2D). *For the same  $C$  as before, for all  $\varepsilon < \min\{1, \mathcal{H}^1(A)^3\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^1(A) \min\{a-1, C\varepsilon^{\frac{2}{3}}\}.$$

*Proof.* This follows immediately from Theorem 3.3.1 and the fact that

$$\Delta\mathcal{E}^{\varepsilon,a,A} = \min_{\chi} \mathcal{E}^{\varepsilon,a,A}[\chi] - W_1(\mu_0, \mu_1) \leq (a-1)W_1(\mu_0, \mu_1) = \mathcal{H}^1(A)(a-1),$$

where we have used  $\mathcal{E}^{\varepsilon,a,A}[\chi] = aW_1(\mu_0, \mu_1)$  for the irrigation pattern  $\chi : \Gamma \times I \rightarrow \mathbb{R}^2$ ,  $\chi(p, t) = (p, t)$ , with reference space  $\Gamma = A$  and reference measure  $P_{\Gamma} = \mathcal{L}^1 \llcorner A$ .  $\square$

*Proof of Theorem 3.3.1.* In the two-dimensional construction we set the height of the top layer of Wasserstein cells to zero,

$$H = 0.$$

Using Lemmas 3.2.2 and 3.2.4 in (3.3.7) we thus obtain

$$\begin{aligned} \Delta\mathcal{E}^{\varepsilon,a,A} &\leq 2 \left[ \sum_{k=1}^K N_k \left( 2(f_k + \varepsilon) \sqrt{\frac{w_k^2}{16} + h_k^2} - w_k h_k \right) + N_{K+1}(f_{K+1} + \varepsilon) w_{K+1} \right] \\ &= \mathcal{H}^1(A) 2 \left[ \sum_{k=1}^K h_k \left( \left( 1 + \frac{2\varepsilon}{w_k} \right) \sqrt{1 + \frac{w_k^2}{16h_k^2}} - 1 \right) + \left( \frac{w_{K+1}}{2} + \varepsilon \right) \right] \\ &\leq \mathcal{H}^1(A) 2 \left[ \sum_{k=1}^K \left( \frac{w_k^2}{32h_k} + \varepsilon \frac{h_k}{w_k} \left( 2 + \frac{w_k^2}{16h_k^2} \right) \right) + \left( \frac{w_{K+1}}{2} + \varepsilon \right) \right], \end{aligned}$$

where we have used  $\sqrt{1+z} \leq 1 + \frac{z}{2}$ . The parameter  $a$  does not occur explicitly in this estimate so that we expect

$$\gamma = 0, \quad \delta = 0$$

to be optimal. Now intuitively, the mass flux in each elementary cell should have a vertical component at most comparable to the horizontal one, so we will continue layering only as long as  $2cw_k \leq h_k$  or equivalently, assuming  $\alpha > 1$ , as long as  $2^{2-k}w_1 \geq \varepsilon^{\frac{\beta}{1-\alpha}}$ . Thus we set

$$K = \left\lfloor 2 + \log_2(w_1 \varepsilon^{\frac{\beta}{\alpha-1}}) \right\rfloor,$$

where the  $\lfloor \cdot \rfloor$  denotes the integer part (note that we will have to verify a posteriori that  $K \geq 1$ ). As a result,  $\frac{w_k^2}{16h_k^2} \leq \frac{1}{64c^2}$  for all  $k \leq K$  as well as  $w_{K+1} < \varepsilon^{\frac{\beta}{1-\alpha}}$ , and the excess energy is estimated as

$$\begin{aligned} \Delta\mathcal{E}^{\varepsilon,a,A} &\leq \mathcal{H}^1(A) \left[ \sum_{k=1}^K \left( \frac{w_k^2}{16h_k} + \varepsilon \frac{h_k}{w_k} \left( 4 + \frac{1}{32c^2} \right) \right) + \left( \varepsilon^{\frac{\beta}{1-\alpha}} + 2\varepsilon \right) \right] \\ &= \mathcal{H}^1(A) \left[ \sum_{k=1}^K \left( \frac{1}{16c} w_k^{2-\alpha} \varepsilon^{-\beta} + \left( 4c + \frac{1}{32c} \right) w_k^{\alpha-1} \varepsilon^{1+\beta} \right) + \left( \varepsilon^{\frac{\beta}{1-\alpha}} + 2\varepsilon \right) \right] \\ &= \mathcal{H}^1(A) \left[ \sqrt{2} \left( \frac{3}{32c} + 4c \right) w_1^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} \sum_{k=1}^K 2^{-\frac{k}{2}} + 3\varepsilon \right] \\ &\leq \mathcal{H}^1(A) \left[ \frac{2}{1-\sqrt{2}} \left( \frac{3}{32c} + 4c \right) w_1^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + 3\varepsilon \right], \end{aligned}$$

where in the third step we chose the optimal exponents

$$\alpha = \frac{3}{2}, \quad \beta = -\frac{1}{2}$$

to minimise the summands and in the last step we exploited the boundedness of the geometric series. Note that for this choice  $w_1$  turns into  $w_1 = \tilde{c}\varepsilon^{\frac{1}{3}}$  and the feasibility conditions (3.3.6) are indeed fulfilled, while constraint (3.3.4) becomes  $\varepsilon \leq \mathcal{H}^1(A)^3$ . Thus we finally arrive at  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq C\mathcal{H}^1(A)\varepsilon^{\frac{2}{3}}$  for some  $C > 0$ , the desired result.  $\square$

### 3.3.2 Upper bound in 3D

In three dimensions there are three different parameter regimes as indicated by Theorems 3.3.3 and 3.3.4 as well as Corollary 3.3.5.

**Theorem 3.3.3** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in 3D for small  $a$ ). *Using the notation from Section 2.2, there exists a constant  $C \geq 0$  such that for all  $a \leq 2$  and  $\varepsilon \leq \min\{\frac{a^2(a-1)^2}{64}, \mathcal{H}^2(A)^2\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^2(A)\sqrt{\varepsilon}C \left[ \left| \log_2 \left( \frac{\sqrt{\varepsilon}}{a(a-1)} \right) \right| + a(a-1) + 1 \right].$$

**Theorem 3.3.4** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in 3D for large  $a$ ). *There exists a constant  $C \geq 0$  such that for all  $\varepsilon \leq \min\{1, \mathcal{H}^2(A)^2\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^2(A)\sqrt{\varepsilon}C \left[ |\log_2(\sqrt{\varepsilon})| + \sqrt{a} + 1 \right].$$

Theorems 3.3.3 and 3.3.4 as well as a third parameter regime can be summarised in the following corollary.

**Corollary 3.3.5** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in 3D). *There exists  $C > 0$  independent of  $a, \varepsilon, A$ , such that for all  $\varepsilon \leq \min(1, \mathcal{H}^2(A)^2)$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^2(A) \min \left\{ a-1, C\sqrt{\varepsilon}(|\log_2 \frac{\sqrt{\varepsilon}}{a-1}| + \sqrt{a}) \right\}.$$

*Proof.* Again, as in Corollary 3.3.2 we have

$$\Delta\mathcal{E}^{\varepsilon,a,A} = \min_{\chi} \mathcal{E}^{\varepsilon,a,A}[\chi] - W_1(\mu_0, \mu_1) \leq (a-1)W_1(\mu_0, \mu_1) = \mathcal{H}^2(A)(a-1)$$

due to  $\mathcal{E}^{\varepsilon,a,A}[\chi] = aW_1(\mu_0, \mu_1)$  for the irrigation pattern  $\chi : \Gamma \times I \rightarrow \mathbb{R}^2$ ,  $\chi(p, t) = (p, t)$ , with reference space  $\Gamma = A$  and reference measure  $P_{\Gamma} = \mathcal{L}^2 \llcorner A$ .

Now let  $\varepsilon \leq \min\{1, \mathcal{H}^2(A)^2\}$ . Without loss of generality we may assume the constant  $C$  in Theorems 3.3.3 and 3.3.4 to be the same and to be greater than 2. Let us abbreviate

$$\begin{aligned} T_1 &= \sqrt{\varepsilon} \left[ \left| \log_2 \left( \frac{\sqrt{\varepsilon}}{a(a-1)} \right) \right| + a(a-1) + 1 \right], \\ T_2 &= \sqrt{\varepsilon} \left[ |\log_2(\sqrt{\varepsilon})| + \sqrt{a} + 1 \right]. \end{aligned}$$

If  $a > 2$ , then it can easily be checked that  $CT_1 > CT_2$ . Likewise, if  $\varepsilon > \frac{a^2(a-1)^2}{64}$  it is straightforward to see  $CT_1 > a-1$ . Thus,  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^2(A) \min\{a-1, CT_1, CT_2\}$  does not only hold under the conditions of Theorem 3.3.3, but even independently of  $a$  and  $\varepsilon$ . Furthermore, we trivially have  $CT_1 \leq 4C\sqrt{\varepsilon}^{\frac{1}{4}}(|\log_2 \frac{\sqrt{\varepsilon}}{a(a-1)}| + a(a-1) + 1)$  and  $CT_2 \leq 4C\sqrt{\varepsilon}(|\log_2 \sqrt{\varepsilon}| + \sqrt{a})$  due to  $a > 1$ , so it remains to show

$$\min \left\{ \frac{1}{4} \left( \left| \log_2 \frac{\sqrt{\varepsilon}}{a(a-1)} \right| + a(a-1) + 1 \right), |\log_2 \sqrt{\varepsilon}| + \sqrt{a} \right\} \leq |\log_2 \frac{\sqrt{\varepsilon}}{a-1}| + \sqrt{a}.$$

For  $a \geq 2$  we have  $|\log_2 \sqrt{\varepsilon}| \leq |\log_2 \frac{\sqrt{\varepsilon}}{a-1}|$  so that the above inequality holds true. For  $a \in (1, 2)$  on the other hand we have

$$\left| \log_2 \frac{\sqrt{\varepsilon}}{a(a-1)} \right| + a(a-1) + 1 \leq \left| \log_2 \frac{\sqrt{\varepsilon}}{a-1} \right| + \log_2 a + a(a-1) + 1 \leq 4 \left| \log_2 \frac{\sqrt{\varepsilon}}{a-1} \right| + 4\sqrt{a}$$

so that the inequality again holds true, which proves the desired inequality in all cases.  $\square$



*Proof of Theorem 3.3.3.* Again we just appeal to Lemmas 3.2.2 and 3.2.4 in (3.3.7) to obtain

$$\begin{aligned}
\Delta \mathcal{E}^{\varepsilon, a, A} &\leq 2 \left[ \sum_{k=1}^K N_k^2 \left( 4(f_k + \varepsilon) \sqrt{\frac{w_k^2}{8} + h_k^2} - w_k^2 h_k \right) + N_{K+1}^2 \left( 4a f_{K+1} \sqrt{\frac{w_{K+1}^2}{8} + H^2} - w_{K+1}^2 H \right) \right] \\
&= \mathcal{H}^2(A) 2 \left[ \sum_{k=1}^K h_k \left( \left( 1 + \frac{4\varepsilon}{w_k} \right) \sqrt{1 + \frac{w_k^2}{8h_k^2}} - 1 \right) + H \left( a \sqrt{1 + \frac{w_{K+1}^2}{8H^2}} - 1 \right) \right] \\
&\leq \mathcal{H}^2(A) 2 \left[ \sum_{k=1}^K \left( \frac{w_k^2}{16h_k} + \varepsilon \frac{h_k}{w_k} \left( 4 + \frac{w_k^2}{4h_k^2} \right) \right) + \left( H(a-1) + a \frac{w_{K+1}^2}{16H} \right) \right]
\end{aligned}$$

using  $\sqrt{1+z} \leq 1 + \frac{z}{2}$ . We now choose the minimising

$$H = \sqrt{\frac{a}{a-1}} \frac{w_{K+1}}{4}$$

(note that we still need to verify  $H \leq \frac{1}{4}$ ), thus arriving at

$$\begin{aligned}
\Delta \mathcal{E}^{\varepsilon, a, A} &\leq \mathcal{H}^2(A) 2 \left[ \sum_{k=1}^K \left( \frac{w_k^2}{16h_k} + \varepsilon \frac{h_k}{w_k} \left( 4 + \frac{w_k^2}{4h_k^2} \right) \right) + \sqrt{a(a-1)} \frac{w_{K+1}}{2} \right] \\
&= \mathcal{H}^2(A) \left[ \sum_{k=1}^K \left( \frac{1}{8c} w_k^{2-\alpha} \varepsilon^{-\beta} + w_k^{\alpha-2} \varepsilon^{1+\beta} \left( 8c + \frac{1}{2c} w_k^{2-2\alpha} \varepsilon^{-2\beta} \right) \right) + \sqrt{a(a-1)} w_{K+1} \right] \\
&= \mathcal{H}^2(A) \left[ \sqrt{\varepsilon} \sum_{k=1}^K \left( \frac{1}{8c} + 8c + \frac{1}{2c} w_k^{-2} \varepsilon \right) + \sqrt{a(a-1)} w_{K+1} \right],
\end{aligned}$$

where we have chosen

$$\gamma = 0, \quad \delta = 0, \quad \alpha = 2, \quad \beta = -\frac{1}{2}$$

to balance all terms optimally. Recall that due to (3.3.3) and (3.3.4) this leads to  $w_1 = \tilde{c} \varepsilon^{\frac{1}{4}}$  as well as the constraint  $\varepsilon \leq \mathcal{H}^2(A)^2$ .

Note that we still have to specify  $K$ . The intuition here is the following. Assuming that  $w_k^2$  stays larger than  $\varepsilon$ , the sum evaluates to a constant times  $K$ . One might expect  $K$  to increase at most logarithmically in  $\varepsilon$  and  $a-1$ , which is not much worse than having a constant upper bound for  $K$ . Thus, the second term may become as large as  $\sqrt{\varepsilon}$  without polluting the energy estimate. Therefore, we will choose  $K$  such that  $w_{K+1} \sim \sqrt{\frac{\varepsilon}{a(a-1)}}$ , that is,

$$K = \left\lfloor \log_2 \frac{\sqrt{a(a-1)}}{\varepsilon^{1/4}} \right\rfloor, \quad \text{and thus } w_{K+1} = \frac{\tilde{c}}{\hat{c}} \sqrt{\frac{\varepsilon}{a(a-1)}}, \quad H = \frac{\tilde{c} \sqrt{\varepsilon}}{4 \hat{c} (a-1)},$$

where again  $\lfloor \cdot \rfloor$  denotes the integer part and where  $\hat{c} = \frac{\varepsilon^{1/4}}{\sqrt{a(a-1)}} 2^K \in (\frac{1}{2}, 1]$ . Note that our assumptions on  $\varepsilon$  ensure  $K \geq 1$  and  $H \leq \frac{1}{8}$  so that condition (3.3.6) is fulfilled.

Inserting our parameter choices in the energy estimate, we have

$$\begin{aligned}
\Delta \mathcal{E}^{\varepsilon, a, A} &\leq \mathcal{H}^2(A) \sqrt{\varepsilon} \left[ \left( \frac{1}{8c} + 8c \right) K + \sum_{k=1}^K \frac{1}{2c} w_k^{-2} \varepsilon + \frac{\tilde{c}}{\hat{c}} \right] \\
&\leq \mathcal{H}^2(A) \sqrt{\varepsilon} \left[ \left( \frac{4}{45} + \frac{45}{4} \right) \frac{1}{2} \left| \log_2 \frac{\varepsilon^{1/2}}{a(a-1)} \right| + \frac{2}{c \tilde{c}^2} \sqrt{\varepsilon} \sum_{k=1}^K 4^{-k} + \frac{\tilde{c}}{\hat{c}} \right] \\
&\leq \mathcal{H}^2(A) \sqrt{\varepsilon} C \left[ \left| \log_2 \frac{\varepsilon^{1/2}}{a(a-1)} \right| + a(a-1) + 1 \right]
\end{aligned}$$

for some constant  $C > 0$ , where we have used the condition on  $\varepsilon$  and the bounds on  $c, \tilde{c}, \hat{c}$ .  $\square$

*Proof of Theorem 3.3.4.* This construction is intended for large values of  $a$  so that unlike in the previous construction, the mass is not discharged from the network at distance  $H$  from the top boundary so as to move the last piece on its own. Instead, the network of pipes goes all the way to the top within  $K_S$  layers and then even continues branching inside the boundary plane before discharging the mass.

We use the same definitions (3.3.1) as before, only in layers  $K_S + 1$  through  $K$  we use

$$h_k = 0, \quad k = K_S + 1, \dots, K, \quad H = 0.$$

Applying Lemmas 3.2.2 and 3.2.4 in (3.3.7) we now obtain the energy estimate

$$\begin{aligned} \Delta \mathcal{E}^{\varepsilon, a, A} &\leq 2 \left[ \sum_{k=1}^{K_S} N_k^2 \left( 4(f_k + \varepsilon) \sqrt{\frac{w_k^2}{8} + h_k^2} - w_k^2 h_k \right) + \sum_{k=K_S+1}^K N_k^2 4(f_k + \varepsilon) \frac{w_k}{\sqrt{8}} + N_{K+1}^2 4a f_{K+1} \frac{w_{K+1}}{\sqrt{8}} \right] \\ &\leq \mathcal{H}^2(A) 2 \left[ \sum_{k=1}^{K_S} \left( \frac{w_k^2}{16h_k} + \varepsilon \frac{h_k}{w_k^2} \left( 4 + \frac{w_k^2}{4h_k^2} \right) \right) + \sum_{k=K_S+1}^K \left( \frac{w_k}{\sqrt{8}} + \frac{4\varepsilon}{\sqrt{8}w_k} \right) + a \frac{w_{K+1}}{\sqrt{8}} \right] \\ &= \mathcal{H}^2(A) \left[ \sqrt{\varepsilon} \sum_{k=1}^{K_S} \left( \frac{1}{8c} + 8c + \frac{1}{2c} w_k^{-2} \varepsilon \right) + \sum_{k=K_S+1}^K \left( \frac{w_k}{\sqrt{2}} + \frac{4\varepsilon}{\sqrt{2}w_k} \right) + a \frac{w_{K+1}}{\sqrt{2}} \right] \end{aligned}$$

for  $\alpha = 2$ ,  $\beta = -\frac{1}{2}$ ,  $\gamma = 0$ , and  $\delta = 0$  as before. Note that due to  $h_k = 0$  for  $k > K_S$  the constant  $c$  is here defined by (3.3.5) with  $H = 0$  and  $K$  replaced by  $K_S$ , which does not enlarge its range, though, as long as we ensure  $K_S \geq 1$ . Again we still need to determine  $K_S$  and  $K$ , and here the heuristics are as follows. Writing the energy contribution from the top two layers as

$$\frac{w_K}{\sqrt{2}} + \frac{4\varepsilon}{\sqrt{2}w_K} + a \frac{w_K}{2\sqrt{2}},$$

we see that the first term can be neglected in view of the third. The minimising  $w_K$  thus satisfies  $w_K \sim \sqrt{\varepsilon/a}$ . Likewise, for the summands of the second sum we would prefer to have the minimising value  $w_k \sim 2\sqrt{\varepsilon}$ . This is not possible for all summands, of course, since the  $w_k$  have to decrease down to  $w_K \sim \sqrt{\varepsilon/a}$ , but at least we can require  $w_k \in [\sqrt{\varepsilon}, \sqrt{\varepsilon/a}]$  for all  $k > K_S$  and  $w_{K_S} \sim \sqrt{\varepsilon}$ . Thus, recalling  $w_1 = \tilde{c}\varepsilon^{\frac{1}{4}}$  from (3.3.3) we need to choose

$$K_S = \left\lfloor \log_2(4\varepsilon^{-\frac{1}{4}}) \right\rfloor, \quad K = K_S + \left\lfloor \log_2 \sqrt{a} \right\rfloor$$

such that (using  $\hat{c} = \varepsilon^{1/4} 2^{K_S} \in (\frac{1}{2}, 1]$  and  $\bar{c} = 2^K / (2^{K_S} \sqrt{a}) \in (\frac{1}{2}, 1]$ ) we indeed have

$$w_{K_S} = \frac{\tilde{c}}{4\hat{c}} \sqrt{\varepsilon}, \quad w_K = \frac{\tilde{c}}{4\hat{c}\bar{c}} \sqrt{\frac{\varepsilon}{a}}.$$

This way the energy estimate turns into

$$\begin{aligned} \Delta \mathcal{E}^{\varepsilon, a, A} &\leq \mathcal{H}^2(A) \left[ \left( \frac{1}{8c} + 8c \right) K \sqrt{\varepsilon} + \sum_{k=1}^{K_S} \frac{1}{2c} w_k^{-2} \varepsilon^{\frac{3}{2}} + \sum_{k=1}^{K-K_S} \left( \frac{w_{K_S}}{2^k 4\sqrt{2}} + \frac{16\varepsilon 2^k}{\sqrt{2}w_{K_S}} \right) + \frac{\tilde{c}}{8\sqrt{2}\hat{c}\bar{c}} \sqrt{a\varepsilon} \right] \\ &= \mathcal{H}^2(A) \sqrt{\varepsilon} \left[ \left( \frac{1}{8c} + 8c \right) K + \frac{2}{c\bar{c}^2} \sqrt{\varepsilon} \sum_{k=1}^{K_S} 4^{-k} + \sum_{k=1}^{K-K_S} \left( \frac{\tilde{c}}{16\hat{c}2^k\sqrt{2}} + \frac{64\hat{c}2^k}{\sqrt{2}\tilde{c}} \right) + \frac{\tilde{c}}{8\sqrt{2}\hat{c}\bar{c}} \sqrt{a} \right] \\ &\leq \mathcal{H}^2(A) \sqrt{\varepsilon} \frac{C}{2} \left[ K + \sqrt{\varepsilon} + 2^{K-K_S} + \sqrt{a} \right] \\ &\leq \mathcal{H}^2(A) \sqrt{\varepsilon} C \left[ \left\lfloor \log_2 \sqrt{\varepsilon} \right\rfloor + 1 + \sqrt{a} \right], \end{aligned}$$

where all occurring constants have been subsumed in a large enough constant  $C > 0$ . □

### 3.3.3 Upper bound in $nD$ for $n > 3$

The following two theorems treat the regimes of small  $a$  and of large  $a$ , respectively.

**Theorem 3.3.6** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in  $nD$  for small  $a$ ). *Using the notation from Section 2.2, there exists a constant  $C \geq 0$  such that for all  $a \leq 2$ ,  $\varepsilon \leq \min\{1, \frac{\sqrt{a(a-1)}^{n+1}}{\sqrt{2^{n^2-1}}}, \sqrt{\frac{1}{(2(n-1))^{n-1}} \frac{(a-1)^{n+1}}{a^{n-3}}}, \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a(a-1)}^{\frac{n-3}{n-1}}.$$

**Theorem 3.3.7** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in  $nD$  for large  $a$ ). *There exists a constant  $C \geq 0$  such that for all  $a \geq 2$ ,  $\varepsilon \leq \min\{1, \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a(a-1)}^{\frac{n-2}{n-1}}.$$

Both theorems can be summarised in the following corollary.

**Corollary 3.3.8** (Upper bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  in  $nD$ ). *There exists a constant  $C \geq 0$  such that for all  $\varepsilon \leq \min\{1, \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}\}$  we have*

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A) \min\left\{a-1, C \varepsilon^{\frac{1}{n-1}} \sqrt{a} \sqrt{a-1}^{\frac{n-3}{n-1}}\right\}.$$

*Proof.* The inequality  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A)(a-1)$  follows as for the two- and three-dimensional case. Now let  $a \geq 1$  and  $\varepsilon \leq \min\{1, \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}\}$  be given. Without loss of generality we may assume the constants  $C$  in Theorems 3.3.6 and 3.3.7 to be identical and to be larger than  $\tilde{C}^{\frac{1}{n-1}}$  for  $\tilde{C} = \min\{\sqrt{2^{1-n^2}}, \sqrt{8(n-1)^{1-n}}\}$ . If  $a \geq 2$ , Theorem 3.3.7 implies

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a(a-1)}^{\frac{n-2}{n-1}} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a} \sqrt{a-1}^{\frac{n-3}{n-1}}$$

so that we are done. If  $a < 2$  and  $\varepsilon$  satisfies the conditions in Theorem 3.3.6, then this theorem implies

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a(a-1)}^{\frac{n-3}{n-1}} \leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} \sqrt{a} \sqrt{a-1}^{\frac{n-3}{n-1}},$$

again finishing the proof. Finally, if  $a < 2$ , but the conditions on  $\varepsilon$  from Theorem 3.3.6 are violated, then in particular we have  $\varepsilon > \tilde{C} \sqrt{a(a-1)}^{n+1}$ . However, this implies  $C \varepsilon^{\frac{1}{n-1}} \sqrt{a} \sqrt{a-1}^{\frac{n-3}{n-1}} \geq (a-1)$  so that the inequality to be proven reduces to  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A)(a-1)$ , which concludes the proof.  $\square$

*Proof of Theorem 3.3.6.* In analogy to the proof of Theorem 3.3.3 we obtain

$$\begin{aligned} \Delta\mathcal{E}^{\varepsilon,a,A} &\leq 2 \left[ \sum_{k=1}^K N_k^{n-1} \left( 2^{n-1} (f_k + \varepsilon) \sqrt{\frac{n-1}{16} w_k^2 + h_k^2} - w_k^{n-1} h_k \right) \right. \\ &\quad \left. + N_{K+1}^{n-1} \left( 2^{n-1} a f_{K+1} \sqrt{\frac{n-1}{16} w_{K+1}^2 + H^2} - w_{K+1}^{n-1} H \right) \right] \\ &\leq \mathcal{H}^{n-1}(A) 2 \left[ \sum_{k=1}^K \left( \frac{(n-1)w_k^2}{32h_k} + \varepsilon \frac{h_k}{w_k^{n-1}} 2^{n-1} \left( 1 + \frac{(n-1)w_k^2}{32h_k^2} \right) \right) + \left( H(a-1) + a \frac{(n-1)w_{K+1}^2}{32H} \right) \right] \\ &= \mathcal{H}^{n-1}(A) 2 \left[ \sum_{k=1}^K \left( \frac{(n-1)w_k^{2-\alpha}}{32c} a^{-\gamma} (a-1)^{-\delta} \varepsilon^{-\beta} + 2^{n-1} c w_k^{\alpha-n+1} a^{\gamma} (a-1)^{\delta} \varepsilon^{1+\beta} \left( 1 + \frac{(n-1)w_k^2}{32h_k^2} \right) \right) \right. \\ &\quad \left. + \sqrt{\frac{n-1}{8}} a(a-1) w_{K+1} \right], \end{aligned} \tag{3.3.8}$$

where we have chosen the minimising

$$H = \sqrt{\frac{n-1}{32} \frac{a}{a-1}} w_{K+1}$$

(for which we still have to verify  $H \leq \frac{1}{4}$ ). Next we choose the exponents

$$\alpha = \frac{n+1}{2}, \quad \beta = -\frac{1}{2}, \quad \gamma = \delta = 0$$

to balance all the terms in the sum, resulting in

$$\Delta\mathcal{E}^{\varepsilon,a,A} \leq \mathcal{H}^{n-1}(A)2 \left[ \sum_{k=1}^K w_k^{\frac{3-n}{2}} \sqrt{\varepsilon} \left( \frac{(n-1)}{32c} + 2^{n-1}c \left( 1 + \frac{(n-1)w_k^2}{32h_k^2} \right) \right) + \sqrt{\frac{n-1}{8}a(a-1)}w_{K+1} \right].$$

Note that by (3.3.3) and (3.3.4) this results in  $w_1 = \tilde{c}^{n+1}\sqrt{\varepsilon}$  as well as the constraint  $\varepsilon \leq \mathcal{H}^{n-1}(A)^{\frac{n+1}{n-1}}$ .

To fix  $K$  the basic idea is as follows. Let us assume  $h_k \geq w_k$  so that the term of the form  $\frac{w_k^2}{h_k^2}$  can be neglected in our estimate. In the above sum the  $w_k$  occur with negative powers so that the sum evaluates to  $\sim \sqrt{\varepsilon}w_{K+1}^{\frac{3-n}{2}}$ . Balancing this with the term of the form  $\sqrt{a(a-1)}w_{K+1}$  we should have  $w_{K+1} \sim \sqrt{n-1} \frac{\varepsilon}{a(a-1)}$ . Thus we choose

$$K = \left\lfloor \log_2 \left( (a(a-1))^{\frac{1}{n-1}} \varepsilon^{-\frac{2}{n^2-1}} \right) \right\rfloor,$$

which due to our assumptions on  $\varepsilon$  is strictly positive and thus admissible (the square brackets again denote the integer part).

With this choice (and using  $\hat{c} = (a(a-1))^{-\frac{1}{n-1}} \varepsilon^{\frac{2}{n^2-1}} 2^K \in (\frac{1}{2}, 1]$ ) we have

$$w_{K+1} = \frac{w_1}{2^K} = \frac{\tilde{c}}{\hat{c}} \left( \frac{\varepsilon}{a(a-1)} \right)^{\frac{1}{n-1}} \quad \text{and} \quad H = \frac{\tilde{c}}{\hat{c}} \sqrt{\frac{n-1}{32} \frac{a}{a-1}} \left( \frac{\varepsilon}{a(a-1)} \right)^{\frac{1}{n-1}},$$

which indeed satisfies  $H \leq \frac{1}{4}$  due to our constraints on  $\varepsilon$ . Thus, all conditions (3.3.6) are fulfilled.

Note that  $\frac{w_k^2}{h_k^2} \leq \frac{w_{K+1}^2}{h_{K+1}^2} = \frac{1}{c^2} \left( \frac{\hat{c}}{\tilde{c}} \right)^{n-1} a(a-1) \leq \frac{2^n}{c^2}$  so that the energy estimate becomes

$$\begin{aligned} \Delta\mathcal{E}^{\varepsilon,a,A} &\leq \mathcal{H}^{n-1}(A)2 \left[ \sum_{k=1}^K w_k^{\frac{3-n}{2}} \sqrt{\varepsilon} \left( \frac{(n-1)}{32c} + 2^{n-1}c \left( 1 + \frac{(n-1)2^{n-5}}{c^2} \right) \right) + \sqrt{\frac{n-1}{8}a(a-1)}w_{K+1} \right] \\ &\leq \mathcal{H}^{n-1}(A) \frac{C}{2} \left[ w_{K+1}^{\frac{3-n}{2}} \sqrt{\varepsilon} + \sqrt{a(a-1)}w_{K+1} \right] \\ &\leq \mathcal{H}^{n-1}(A) C \varepsilon^{\frac{1}{n-1}} (a(a-1))^{\frac{n-3}{2(n-1)}} \end{aligned}$$

for some  $C > 0$ . □

*Proof of Theorem 3.3.7.* For large  $a$  the previous energy estimate is no longer correct since it is based on  $h_k \gtrsim w_k$ , which the construction will not always satisfy for large  $a$ . This indicates that the term involving  $\frac{w_k^2}{h_k^2}$  in (3.3.8) will become dominant. As before, the summands in (3.3.8) will be negligible except for the  $K^{\text{th}}$  term so that the energy behaves up to constant factors like

$$\Delta\mathcal{E}^{\varepsilon,a,A} \sim \mathcal{H}^{n-1}(A) \left[ \left( w_K^{2-\alpha} a^{-\gamma} (a-1)^{-\delta} \varepsilon^{-\beta} + w_K^{\alpha-n+1} a^{\gamma} (a-1)^{\delta} \varepsilon^{1+\beta} \left( 1 + \frac{w_K^2}{h_K^2} \right) \right) + \sqrt{a(a-1)}w_K \right].$$

Let us assume the first term to be negligible. If we choose optimal parameters, the remaining three terms should balance, that is, we should have  $1 \sim \frac{w_K^2}{h_K^2}$  as well as  $w_K^{\alpha-n+1} a^{\gamma} (a-1)^{\delta} \varepsilon^{1+\beta} \sim \sqrt{a(a-1)}w_K$ . The former results in  $w_K \sim h_K$  or equivalently

$$w_K \sim a^{\frac{\gamma}{1-\alpha}} (a-1)^{\frac{\delta}{1-\alpha}} \varepsilon^{\frac{\beta}{1-\alpha}}$$

while the latter is equivalent to

$$w_K \sim \left( a^{\frac{1}{2}-\gamma} (a-1)^{\frac{1}{2}-\delta} \varepsilon^{-1-\beta} \right)^{\frac{1}{\alpha-n}}.$$

Equating both scalings and picking  $\alpha$  as in the previous proof we obtain the exponents

$$\alpha = \frac{n+1}{2}, \quad \beta = -\frac{1}{2}, \quad \gamma = \delta = \frac{1}{4}.$$

The above scaling of  $w_K$  together with  $w_K = w_1 2^{1-K} = \tilde{c} \left( \frac{\varepsilon}{\sqrt{a(a-1)}} \right)^{\frac{1}{n+1}} 2^{1-K}$  now suggests to choose

$$K = \left\lfloor \log_2 \left( 2\tilde{c}\sqrt{n-1}(a(a-1))^{\frac{1}{n^2-1}} \varepsilon^{-\frac{2}{n^2-1}} \right) \right\rfloor \geq 1.$$

For some constant  $\hat{c} \in (\frac{1}{2}, 1]$  this then results in

$$w_{K+1} = \frac{1}{2\hat{c}\sqrt{n-1}} \left( \frac{\varepsilon}{\sqrt{a(a-1)}} \right)^{\frac{1}{n-1}} \text{ and } H = \frac{1}{\hat{c}\sqrt{128}} \sqrt{\frac{a}{a-1}} \left( \frac{\varepsilon}{\sqrt{a(a-1)}} \right)^{\frac{1}{n-1}} \leq \frac{1}{4},$$

satisfying (3.3.6). Finally, the energy estimate (3.3.8) turns into

$$\Delta \mathcal{E}^{\varepsilon, a, A} \leq \mathcal{H}^{n-1}(A) C \sqrt{a(a-1)}^{\frac{n-2}{n-1}} \varepsilon^{\frac{1}{n-1}}$$

for some constant  $C > 0$ . □

*Remark 3.3.9.* It is a straightforward exercise to show that in the proof of Theorem 3.3.6 we could have equally well used any exponent  $\alpha \in (2, n-1)$  as well as  $\beta = \frac{1-\alpha}{n-1}$ ,  $\gamma = \delta = \frac{\alpha-1}{n-1} - \frac{1}{2}$ , and  $K \sim \log_2 \left( (a(a-1))^{\frac{n+1}{2\alpha(n-1)}} \varepsilon^{-\frac{1}{\alpha(n-1)}} \right)$ . Likewise, in the proof of Theorem 3.3.7 one could also choose  $\alpha \in (2, n-1)$ ,  $\beta = \frac{1-\alpha}{n-1}$ ,  $\gamma = \delta = -\frac{\beta}{2}$ , and  $K \sim \log_2 \left( (a(a-1))^{\frac{1}{2\alpha(n-1)}} \varepsilon^{-\frac{1}{\alpha(n-1)}} \right)$ . This tells us that the energy scaling does not give highly precise information about how optimal geometries and irrigation patterns look like, since there are multiple rather different geometries with the same energy scaling.

### 3.4 Upper bound for optimal branched transport energy $\mathcal{M}^{\varepsilon, A}$

Here the construction of an irrigation pattern with the desired energy scaling is similar to the urban planning case, however, this time the mass cannot travel outside the pipe network. This implies that no Wasserstein cells can be used and instead the pipe network has to refine infinitely. Again we will assume  $A = [0, \ell]^{n-1}$  without loss of generality.

**Theorem 3.4.1** (Upper bound on  $\Delta \mathcal{M}^{\varepsilon, A}$ ). *Using the notation from Section 2.3, there exists  $C > 0$  independent of  $\varepsilon, A$  such that for all  $\varepsilon < \min\{\frac{1}{2(n-1)}, \mathcal{H}^{n-1}(A)^{\frac{2}{n-1}}\}$  we have*

$$\Delta \mathcal{M}^{\varepsilon, A} \leq \mathcal{H}^{n-1}(A) C \varepsilon |\log \varepsilon|.$$

*Proof.* Again, the width, number, and flux of elementary cells in the  $k^{\text{th}}$  layer are specified as

$$w_k = 2^{1-k} w_1, \quad N_k = \frac{\ell}{w_k}, \quad f_k = \left( \frac{w_k}{2} \right)^{n-1}.$$

Furthermore, letting  $i \in \{1, \dots, N_k\}^{n-1}$  a multiindex, the height of an elementary cell and the  $i^{\text{th}}$  base point shall be

$$h_k = \begin{cases} c w_k^\alpha \varepsilon^\beta, & k \leq K, \\ 0, & k > K, \end{cases} \quad x_{k,i} = \left( (i_1 - \frac{1}{2}) w_k, \dots, (i_{n-1} - \frac{1}{2}) w_k, \frac{1}{2} + \sum_{j=1}^{k-1} h_j \right)$$

for some constants  $c, \alpha, \beta \in \mathbb{R}$ ,  $K \in \mathbb{N}$  to be determined. As in urban planning, the irrigation pattern of the  $k^{\text{th}}$  layer is given by

$$\chi_k = \coprod_{i \in \{1, \dots, N_k\}^{n-1}} \chi_{x_{k,i}, w_k, h_k, f_k}^E,$$

and these patterns are again compatible to be composed in series according to

$$\chi_u = ((\dots((\chi_1 \circ_{\text{Id}} \chi_2) \circ_{\text{Id}} \dots \circ_{\text{Id}} \chi_K) \circ_{\text{Id}} \chi_{K+1}) \circ_{\text{Id}} \chi_{K+2}) \circ_{\text{Id}} \dots$$

In contrast to urban planning this is an infinite series of patterns, but it is straightforward to see that it is actually well-defined (using, for example, that the lengths of the fibres converge). As in the urban planning case, the full irrigation pattern is then defined as

$$\chi = \chi_l \circ_{\text{Id}} \chi_u,$$

where  $\chi_l$  describes the lower half and is constructed analogously to  $\chi_u$ .

Since the total height of the construction is 1 we have

$$1 = 2 \sum_{k=1}^{\infty} h_k = 2 \sum_{k=1}^K c w_k^{\alpha} \varepsilon^{\beta} = 2c(2w_1)^{\alpha} \varepsilon^{\beta} \sum_{k=1}^K 2^{-k\alpha} = 2c w_1^{\alpha} \varepsilon^{\beta} \frac{1 - 2^{-K\alpha}}{1 - 2^{-\alpha}},$$

which implies that we should choose

$$w_1 = \tilde{c} \varepsilon^{-\frac{\beta}{\alpha}}, \quad c = \frac{1}{2\tilde{c}^{\alpha}} \frac{1 - 2^{-\alpha}}{1 - 2^{-K\alpha}} \in [\frac{1 - 2^{-\alpha}}{2}, 2^{\alpha-1}],$$

where we have used  $K \geq 1$ ,  $\alpha > 0$  (which will be verified later) and where  $\tilde{c} \in (\frac{1}{2}, 1]$  is such that the total geometry width  $\ell$  is a multiple of  $w_1$ . Note that this requires  $\varepsilon \leq \ell^{-\frac{\alpha}{\beta}} = \mathcal{H}^{n-1}(A)^{-\frac{\alpha}{\beta(n-1)}}$ .

Next let us specify  $K$ . The idea behind taking  $h_k = 0$  for all  $k > K$  is to separate all elementary cell layers into those for which a substantial vertical mass flux takes place and those for which the vertical mass flux can be neglected compared to the horizontal one. Thus we should have  $h_k \gtrsim w_k$  for  $k \leq K$  and  $h_K \sim w_K$  or equivalently  $w_K \sim \varepsilon^{\frac{\beta}{1-\alpha}}$ . Due to  $w_K = w_1 2^{1-K} = \tilde{c} \varepsilon^{-\frac{\beta}{\alpha}} 2^{1-K}$  this suggest to fix

$$K = 1 + \left\lceil \log_2 \left( 2\tilde{c} \varepsilon^{\frac{\beta}{\alpha(\alpha-1)}} \right) \right\rceil \text{ and thus } w_K = \frac{1}{2\tilde{c}} \varepsilon^{-\frac{\beta}{\alpha-1}}$$

for some  $\hat{c} \in (\frac{1}{2}, 1]$ , where we have  $K \geq 1$  as long as  $\alpha > 1$  and  $\beta \leq 0$ .

It remains to specify  $\alpha$  and  $\beta$  and to estimate the energy of  $\chi$ . As in the urban planning case we exploit the fact

$$\mathcal{M}^{*,A} = \mathcal{H}^1(A) = 2 \sum_{k=1}^K N_k^{n-1} w_k^{n-1} h_k$$

as well as Remark 3.1.4 and Lemma 3.2.2 to obtain

$$\begin{aligned} \Delta \mathcal{M}^{\varepsilon,A} &= 2M_{\mathbb{P}}^{\varepsilon}[\chi_u] - \mathcal{M}^{*,A} \\ &= 2 \left[ \sum_{k=1}^K \sum_{i \in \{1, \dots, N_k\}^{n-1}} \left( M_{\mathbb{P}}^{\varepsilon}[\chi_{x_k,i,w_k,h_k,f_k}^E] - w_k^{n-1} h_k \right) + \sum_{k=K+1}^{\infty} \sum_{i \in \{1, \dots, N_k\}^{n-1}} M_{\mathbb{P}}^{\varepsilon}[\chi_{x_k,i,w_k,h_k,f_k}^E] \right] \\ &\leq 2 \left[ \sum_{k=1}^K N_k^{n-1} \left( 2^{n-1} \left( \frac{w_k}{2} \right)^{(n-1)(1-\varepsilon)} \sqrt{\frac{n-1}{16} w_k^2 + h_k^2} - w_k^{n-1} h_k \right) \right. \\ &\quad \left. + \sum_{k=K+1}^{\infty} N_k^{n-1} 2^{n-1} \left( \frac{w_k}{2} \right)^{(n-1)(1-\varepsilon)} \frac{\sqrt{n-1}}{4} w_k \right] \\ &\leq 2\mathcal{H}^{n-1}(A) \left[ \sum_{k=1}^K h_k \left( \left( \frac{w_k}{2} \right)^{-\varepsilon(n-1)} \sqrt{\frac{n-1}{16} \frac{w_k^2}{h_k^2} + 1} - 1 \right) + \sum_{k=K+1}^{\infty} \left( \frac{w_k}{2} \right)^{-\varepsilon(n-1)} \frac{\sqrt{n-1}}{4} w_k \right]. \end{aligned}$$

Note that we can estimate

$$\log \left( \left( \frac{w_k}{2} \right)^{-\varepsilon(n-1)} \right) \leq \log \left( \left( \frac{w_K}{2} \right)^{-\varepsilon(n-1)} \right) \leq (n-1)\varepsilon \left( \log(4\hat{c}) + \left| \frac{\beta}{\alpha-1} \right| \log \varepsilon \right) \leq \log 4 + (n-1) \left| \frac{\beta}{\alpha-1} \right|.$$

Abbreviating the right-hand side as  $Z$ , we now employ the inequalities

$$\sqrt{1+z} \leq 1 + \frac{z}{2} \quad \forall z \geq 0 \quad \text{and} \quad e^{\tilde{z}} \leq 1 + \bar{c}\tilde{z} \quad \forall \tilde{z} \in [0, Z]$$

for  $\bar{c} = \frac{\varepsilon^Z - 1}{Z}$ . Using  $z = \frac{n-1}{16} \frac{w_k^2}{h_k^2}$  and  $\tilde{z} = \log\left(\left(\frac{w_k}{2}\right)^{-\varepsilon(n-1)}\right)$ , the energy estimate turns into

$$\begin{aligned} \Delta \mathcal{M}^{\varepsilon, A} &\leq 2\mathcal{H}^{n-1}(A) \left[ \sum_{k=1}^K h_k \left( (1 - \bar{c}\varepsilon(n-1) \log(\frac{w_k}{2})) \left( \frac{n-1}{32} \frac{w_k^2}{h_k^2} + 1 \right) - 1 \right) + \frac{\sqrt{n-1}}{4} \sum_{k=K+1}^{\infty} \left( \frac{w_k}{2} \right)^{1-\varepsilon(n-1)} \right] \\ &\leq 2\mathcal{H}^{n-1}(A) \left[ \bar{c}\varepsilon(n-1) \left| \log \frac{w_K}{2} \right| \sum_{k=1}^K h_k + \left( \frac{n-1}{32} + Z \right) \sum_{k=1}^K \frac{w_k^2}{h_k} + \frac{\sqrt{n-1}}{4} \sum_{k=1}^{\infty} \left( \frac{w_K}{2} \right)^{1-\varepsilon(n-1)} \right] \\ &= 2\mathcal{H}^{n-1}(A) \left[ \frac{\bar{c}\varepsilon(n-1)}{2} \left| \log \frac{1}{4\bar{c}} \varepsilon^{-\frac{\beta}{\alpha-1}} \right| + \left( \frac{n-1}{32} + Z \right) (2w_1)^{2-\alpha} \varepsilon^{-\beta} \sum_{k=1}^K 2^{-k(2-\alpha)} \right. \\ &\quad \left. + \frac{\sqrt{n-1}}{4} \left( \frac{1}{2\bar{c}} \varepsilon^{-\frac{\beta}{\alpha-1}} \right)^{1-\varepsilon(n-1)} \frac{1}{2^{1-\varepsilon(n-1)} - 1} \right]. \end{aligned}$$

Choosing

$$\alpha = 2, \quad \beta = -1,$$

and letting  $\tilde{C}, C > 0$  denote constants independent of  $\varepsilon, A$ , this turns into

$$\Delta \mathcal{M}^{\varepsilon, A} \leq \mathcal{H}^{n-1}(A) \tilde{C} \varepsilon [|\log \varepsilon| + K + \varepsilon^{-(n-1)\varepsilon}] \leq \mathcal{H}^{n-1}(A) C \varepsilon |\log \varepsilon|$$

as desired.  $\square$

## 4 Lower bound based on convex analysis

In this section we prove the lower bounds from Theorems 2.4.1 and 2.4.2. The technique employed is based on the one briefly explained in the introduction and heavily exploits convex duality. In particular, we will several times need the following convex duality bound on the Wasserstein distance.

### 4.1 Bound on Wasserstein distance

**Lemma 4.1.1.** *Given  $s, t \in \mathbb{R}$ , let  $\mu = g\mathcal{H}^{n-1} \llcorner \{x_n = s\}$  for some measurable function  $g : \{x_n = s\} \rightarrow [0, m]$  and  $\nu$  be a non-negative discrete measure with  $\text{spt} \nu \subset \{x_n = t\}$  and  $\mathcal{H}^0(\text{spt} \nu) = N$ . Then there is a positive constant  $C \equiv C(n)$  such that*

$$W_1(\mu, \nu) \geq |t - s| \|\nu\|_{\text{fbm}} + C \|\nu\|_{\text{fbm}} \min\left\{\frac{R^2}{|t-s|}, R\right\}$$

for  $R = \sqrt[n-1]{\frac{\|\nu\|_{\text{fbm}}}{mN\omega_{n-1}}}$  the radius of a single disk if  $\mu$  were a uniform measure of density  $m$  on  $N$  disks of equal size.

*Proof.* If  $\mu$  and  $\nu$  have different mass,  $W_1(\mu, \nu) = \infty$  and there is nothing to prove, so we assume  $\|\mu\|_{\text{fbm}} = \|\nu\|_{\text{fbm}}$ . Also, without loss of generality we shall assume  $s = 0$  and  $t > 0$ .

Let us define the test function

$$\psi(x) = \text{dist}(x, \text{spt} \nu).$$

The dual formulation of the Wasserstein distance (Theorem 1.1.2) yields

$$W_1(\mu, \nu) - t \|\nu\|_{\text{fbm}} \geq \int_{\mathbb{R}^n} \psi \, d(\mu - \nu) - t \|\nu\|_{\text{fbm}} = \int_{\mathbb{R}^n} \text{dist}(x, \text{spt} \nu) - t \, d\mu = \int_0^\infty \mu(\mathbb{R}^n \setminus B_{t+r}(\text{spt} \nu)) \, dr,$$

where we employed the layer cake formula. Note that the integrand is bounded below by

$$\mu(\mathbb{R}^n) - \mu(B_{t+r}(\text{spt} \nu)) \geq \max \left\{ 0, \|\nu\|_{\text{fbm}} - mN\omega_{n-1} \left( \sqrt{(t+r)^2 - t^2} \right)^{n-1} \right\} =: M(r).$$

Introducing the new variable  $\rho = \sqrt{(t+r)^2 - t^2}$  and the maximum radius  $R = \left( \frac{\|\nu\|_{\text{fbm}}}{mN\omega_{n-1}} \right)^{\frac{1}{n-1}}$  we thus obtain the estimate

$$W_1(\mu, \nu) - t \|\nu\|_{\text{fbm}} \geq \int_0^\infty M(r) \, dr = \int_0^R \frac{\|\nu\|_{\text{fbm}} - mN\omega_{n-1} \rho^{n-1}}{\sqrt{1 + (\frac{t}{\rho})^2}} \, d\rho = mN\omega_{n-1} \int_0^R \frac{R^{n-1} - \rho^{n-1}}{\sqrt{1 + (\frac{t}{\rho})^2}} \, d\rho.$$

Now if  $R \leq 2t$ , we have  $\sqrt{1 + (\frac{t}{\rho})^2} \leq \sqrt{5} \frac{t}{\rho}$  for all  $\rho \leq R$  so that

$$W_1(\mu, \nu) - t\|\nu\|_{\text{fbm}} \geq \frac{mN\omega_{n-1}}{\sqrt{5}t} \int_0^R R^{n-1} \rho - \rho^n d\rho = \frac{n-1}{\sqrt{5}(2n+2)} \|\nu\|_{\text{fbm}} \frac{R^2}{t},$$

while for  $R > 2t$  we obtain

$$W_1(\mu, \nu) - t\|\nu\|_{\text{fbm}} \geq mN\omega_{n-1} \int_{R/2}^R \frac{R^{n-1} - \rho^{n-1}}{\sqrt{2}} d\rho = \frac{1 + (n-2)2^{n-1}}{2^n \sqrt{2}n} \|\nu\|_{\text{fbm}} R.$$

□

We are now in a situation to proceed with the lower bound proofs.

## 4.2 Lower bound for $\mathcal{E}^{\varepsilon, a, A}$

In this section we prove the lower bound from Theorem 2.4.1. We will treat the different dimensions separately, as different phenomena occur in different dimensions. In two dimensions, the bulk of the excess energy is contributed from the centre of the domain  $A \times [0, 1]$ , while in dimension greater than three, the dominant part of the excess energy stems from boundary layers at  $x_n = 0$  and  $x_n = 1$ . The three-dimensional situation represents an intermediate case, where the dominant energy contribution is distributed all over  $A \times [0, 1]$ . In all cases the proofs follow variations of the same theme, namely the classical technique from [KM94]. In two dimensions, where the energy is concentrated near the centre, we will bound the excess energy from below using the relaxed energy between  $\mu_0, \mu_1$  and a generic cross-section  $\{x_2 = t\}$  with  $t \sim \frac{1}{2}$ . In three dimensions, where the energy is distributed over different layers, we will employ the same technique, only now considering various cross-sections. In higher dimensions, where the energy is concentrated near the boundary, we will apply the technique to a cross-section near the boundary. The estimates require lower bounds of the urban planning cost in terms of the Wasserstein distance. In detail, for two finite Borel measures  $\mu_+, \mu_-$  of equal mass and an optimal connecting network  $\Sigma^*$  we will use the bound

$$\min_{\chi} \mathcal{E}^{\varepsilon, a, \mu_+, \mu_-}[\chi] = \mathcal{E}^{\varepsilon, a, \mu_+, \mu_-}[\Sigma^*] \geq W_{d_{\Sigma^*}}(\mu_+, \mu_-) \geq W_1(\mu_+, \mu_-). \quad (4.2.1)$$

In the following we denote by  $\Sigma^*$  and  $\chi^*$  the optimal network and irrigation pattern for  $\mathcal{E}^{\varepsilon, a, A}$ , respectively, which are related as detailed in Section 2.2. Furthermore, we use the following type of pattern reparameterisation. For  $s \in (0, 1)$ ,  $p \in \Gamma$  we define the *section crossing time* and the *section crossing point*,

$$t_p(s) = \min \{t \in I : \chi_p^*(t) \in \{x_n = s\}\}, \quad (4.2.2)$$

$$\xi^*(p, s) = \chi_p^*(t_p(s)), \quad (4.2.3)$$

to be the time when particle  $p$  reaches the cross-section  $\{x_n = s\}$  for the first time and its position at that time, respectively. Both the above are well-defined since for almost all  $p \in \Gamma$ ,  $\chi_p^* \in \text{AC}(I)$  and  $\chi_p^*(t) \in \{x_n = t\}$  for  $t = 0, 1$ . If  $\chi_p^*$  visits each cross-section just once,  $\xi_p^* = \xi^*(p, \cdot)$  is indeed a reparameterisation of  $\chi_p^*(I)$ , else it just reparameterises part of  $\chi_p^*(I)$ . In addition we set

$$\Gamma_s = \{p \in \Gamma : \xi_p^*(s) \in \Sigma^*\},$$

$$\Delta F_s = P_{\Gamma}(\Gamma \setminus \Gamma_s),$$

$$\mu_s = \xi^*(\cdot, s)_{\#}(P_{\Gamma} \lfloor \Gamma_s),$$

$$N_s = \mathcal{H}^0(\text{spt} \mu_s)$$

to be the set of particles travelling through the network  $\Sigma^*$ , the amount of particles bypassing the network, the mass distribution inside the network, and the number of network pipes, all at height  $s$ . Note that  $\Gamma_s$  is well-defined up to a null set, that  $N_s$  might possibly be infinite, and that  $\|\mu_s\|_{\text{fbm}} = P_{\Gamma}(\Gamma) - \Delta F_s$ .



Bounds on  $N_s$  and  $\Delta F_s$  can be established via the following estimate. First note that

$$\begin{aligned} \int_I r_{\varepsilon,a}^{\chi^*}(\chi^*(p,t))|\dot{\chi}^*(p,t)|dt &\geq \mathcal{H}^1(\chi_p^*(I)) + \int_{\{\chi_p^*(I) \cap \Sigma^*\}} \frac{\varepsilon}{m_{\chi^*}(x)} d\mathcal{H}^1(x) + \int_{\{\chi_p^*(I) \setminus \Sigma^*\}} (a-1) d\mathcal{H}^1(x) \\ &\geq 1 + \varepsilon \int_{\{\chi_p^*(I) \cap \Sigma^*\}} \frac{1}{m_{\chi^*}(x)} d\mathcal{H}^1(x) + (a-1) \int_{\{\chi_p^*(I) \setminus \Sigma^*\}} d\mathcal{H}^1(x). \end{aligned} \quad (4.2.4)$$

Note also that

$$\begin{aligned} \int_{\Gamma} \int_{\{\chi_p^*(I) \cap \Sigma^*\}} \frac{1}{m_{\chi^*}(x)} d\mathcal{H}^1(x) dP_{\Gamma}(p) &= \int_{\Sigma^*} \int_{\{p \in \Gamma : x \in \chi_p^*(I)\}} \frac{1}{m_{\chi^*}(x)} dP_{\Gamma}(p) d\mathcal{H}^1(x) \\ &= \int_{\Sigma^*} \frac{m_{\chi^*}(x)}{m_{\chi^*}(x)} d\mathcal{H}^1(x) = \mathcal{H}^1(\Sigma^*) \geq \int_0^1 N_s ds. \end{aligned} \quad (4.2.5)$$

We also have

$$\begin{aligned} \int_{\Gamma} \int_{\{\xi_p^*(I) \setminus \Sigma^*\}} d\mathcal{H}^1(x) dP_{\Gamma}(p) &\geq \int_{\Gamma} \int_0^1 \mathbf{1}_{\mathbb{R}^n \setminus \Sigma^*}(\xi_p^*(s)) ds dP_{\Gamma}(p) \\ &= \int_0^1 \int_{\Gamma \setminus \Gamma_s} dP_{\Gamma}(p) ds = \int_0^1 \Delta F_s ds. \end{aligned} \quad (4.2.6)$$

Integrating inequality (4.2.4) over  $\Gamma$  and using inequalities (4.2.5) and (4.2.6) we thus obtain

$$\mathcal{E}^{\varepsilon,a,A}[\chi^*] \geq P_{\Gamma}(\Gamma) + \int_0^1 \varepsilon N_s + (a-1) \Delta F_s ds.$$

#### 4.2.1 Lower bound in 2D

The lower bound derived here actually holds in any dimension, only it is not sharp for  $n > 2$ . Therefore we will keep the more general notation  $n$  for the dimension.

The proof proceeds in the earlier described way: we first state some properties of a generic cross-section and then solve the relaxed problem below and above that cross-section, exploiting the cross-section properties. The cross-section properties are bounds in terms of the excess energy  $\Delta \mathcal{E}^{\varepsilon,a,A}$ .

**Lemma 4.2.1.** *There is a generic  $s \in (0,1)$  such that*

$$N_s \leq \frac{\Delta \mathcal{E}^{\varepsilon,a,A}}{\varepsilon} \quad \text{and} \quad \Delta F_s \leq \frac{\Delta \mathcal{E}^{\varepsilon,a,A}}{a-1}.$$

*Proof.* For a proof by contradiction, assume that almost all cross-sections  $s \in (0,1)$  satisfy  $N_s > \frac{\Delta \mathcal{E}^{\varepsilon,a,A}}{\varepsilon}$  or  $\Delta F_s > \frac{\Delta \mathcal{E}^{\varepsilon,a,A}}{a-1}$ . Then

$$\begin{aligned} \mathcal{E}^{\varepsilon,a,A}[\chi^*] &\geq P_{\Gamma}(\Gamma) + \int_0^1 \varepsilon N_s + (a-1) \Delta F_s ds \\ &= \mathcal{H}^1(A) + \int_0^1 \varepsilon N_s + (a-1) \Delta F_s ds \\ &\geq \mathcal{E}^{*,a,A} + \int_0^1 \max\{\varepsilon N_s, (a-1) \Delta F_s\} ds \\ &> \mathcal{E}^{*,a,A} + \Delta \mathcal{E}^{\varepsilon,a,A}, \end{aligned}$$

the desired contradiction. □

**Proposition 4.2.2.** *For  $\varepsilon \leq 1$  and  $a > 1$  we have*

$$\Delta \mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A) \min\{\varepsilon^{\frac{2}{n+1}}, a-1\}.$$

*Proof.* Let  $s \in (0, 1)$  be the generic cross-section of the previous lemma, and let

$$\tilde{\mu}_0 = \chi^*(\cdot, 0)_{\#}(P_{\Gamma} \llcorner \Gamma_s), \quad (4.2.7)$$

$$\tilde{\mu}_1 = \chi^*(\cdot, 1)_{\#}(P_{\Gamma} \llcorner \Gamma_s) \quad (4.2.8)$$

be the irrigating and the irrigated measure, respectively, of all particles in  $\Gamma_s$ . Abbreviating  $R = \frac{n-1}{N_s \omega_{n-1}} \sqrt{\frac{\|\mu_s\|_{\text{fbm}}}{N_s \omega_{n-1}}}$ , we have (see Remark 3.1.4)

$$\begin{aligned} \mathcal{E}^{\varepsilon, a, A}[\chi^*] &= \mathcal{E}^{\varepsilon, a, \tilde{\mu}_0, \mu_s}[\chi^*|_{\Gamma_s \times [0, s]}] + \mathcal{E}^{\varepsilon, a, \mu_s, \tilde{\mu}_1}[\chi^*|_{\Gamma_s \times [s, 1]}] + \mathcal{E}^{\varepsilon, a, \mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1}[\chi^*|_{(\Gamma \setminus \Gamma_s) \times I}] \\ &\geq W_1(\tilde{\mu}_0, \mu_s) + W_1(\mu_s, \tilde{\mu}_1) + W_1(\mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1) \\ &\geq s\|\mu_s\|_{\text{fbm}} + C\|\mu_s\|_{\text{fbm}} \min\left\{\frac{R^2}{s}, R\right\} + (1-s)\|\mu_s\|_{\text{fbm}} + C\|\mu_s\|_{\text{fbm}} \min\left\{\frac{R^2}{1-s}, R\right\} + \|\mu_0 - \tilde{\mu}_0\|_{\text{fbm}} \\ &= \|\mu_0\|_{\text{fbm}} + CR\|\mu_s\|_{\text{fbm}} \left(\min\left\{\frac{R}{s}, 1\right\} + \min\left\{\frac{R}{1-s}, 1\right\}\right) \\ &\geq \mathcal{E}^{*, a, A} + CR\|\mu_s\|_{\text{fbm}} \min\{2R, 1\}, \end{aligned}$$

where we have employed (4.2.1) and Lemma 4.1.1. Now the two cases  $R > \frac{1}{2}$  and  $R \leq \frac{1}{2}$  are treated separately. For  $R > \frac{1}{2}$ , the above inequality turns into

$$\Delta \mathcal{E}^{\varepsilon, a, A} \geq \frac{C}{2} \|\mu_s\|_{\text{fbm}} = \frac{C}{2} (\mathcal{H}^{n-1}(A) - \Delta F_s) \geq \frac{C}{2} \mathcal{H}^{n-1}(A) - \frac{C}{2(a-1)} \Delta \mathcal{E}^{\varepsilon, a, A}$$

by virtue of Lemma 4.2.1, which implies

$$\Delta \mathcal{E}^{\varepsilon, a, A} \geq \mathcal{H}^{n-1}(A) / \left(\frac{2}{C} + \frac{1}{a-1}\right) \gtrsim \min\{1, a-1\} \mathcal{H}^{n-1}(A).$$

If  $R \leq \frac{1}{2}$ , the above inequality turns into

$$\Delta \mathcal{E}^{\varepsilon, a, A} \geq 2CR^2 \|\mu_s\|_{\text{fbm}} = 2C \|\mu_s\|_{\text{fbm}}^{1 + \frac{2}{n-1}} \left(\frac{1}{N_s \omega_{n-1}}\right)^{\frac{2}{n-1}} \gtrsim \|\mu_s\|_{\text{fbm}}^{1 + \frac{2}{n-1}} \left(\frac{\varepsilon}{\Delta \mathcal{E}^{\varepsilon, a, A}}\right)^{\frac{2}{n-1}},$$

which can be solved for  $\Delta \mathcal{E}^{\varepsilon, a, A}$  to yield

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \|\mu_s\|_{\text{fbm}} \varepsilon^{\frac{2}{n+1}} \geq \mathcal{H}^{n-1}(A) \varepsilon^{\frac{2}{n+1}} - \frac{1}{a-1} \Delta \mathcal{E}^{\varepsilon, a, A} \varepsilon^{\frac{2}{n+1}}$$

by Lemma 4.2.1. This finally implies

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^{n-1}(A) / \left(\varepsilon^{-\frac{2}{n+1}} + \frac{1}{a-1}\right) \gtrsim \mathcal{H}^{n-1}(A) \min\left\{\varepsilon^{\frac{2}{n+1}}, a-1\right\}.$$

□

## 4.2.2 Lower bound in 3D

The previous estimate in three dimensions becomes  $\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A) \min\{\sqrt{\varepsilon}, a-1\}$ . However, the upper bound and its proof suggest that we are missing an additional factor  $|\log \varepsilon|$ , which comes from the fact that—when the optimal network  $\Sigma^*$  refines from the centre  $x_3 = \frac{1}{2}$  to the boundary—every refinement layer contributes the same excess energy. Therefore, it no longer suffices to consider a single cross-section and the excess energy induced by it, but the previous argument has to be refined to take into account multiple cross-sections. The following lemma characterises the different cross-sections.

**Lemma 4.2.3.** *For any  $K > 0$  there are  $s_k \in [2^{-k-1}, 2^{-k}]$ ,  $k = 1, \dots, 2K$ , with  $s_k - s_{k+1} > 2^{-k-2}$  and such that at least half of the  $s_k$  satisfy*

$$N_{s_k} \leq 4 \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{K \varepsilon s_k} \quad \text{and} \quad \Delta F_{s_k} \leq 4 \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{K(a-1)s_k}.$$

*Proof.* Consider the dyadic intervals  $J_k = [2^{-k-1}, 2^{-k}]$ ,  $k = 1, \dots, 2K$ , and denote by  $J \subset \{1, \dots, 2K\}$  the set of indices  $k$  for which  $N_s > 4 \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{K \varepsilon s}$  or  $\Delta F_s > 4 \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{K(a-1)s}$  for no less than half the  $s \in J_k$ . Then we estimate

$$\Delta \mathcal{E}^{\varepsilon, a, A} \geq \sum_{k=1}^{2K} \int_{J_k} \varepsilon N_s + (a-1) \Delta F_s \, ds > \sum_{k \in J} \int_{3 \cdot 2^{-k-2}}^{2^{-k}} 4 \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{K s} \, ds = \frac{|J|}{K} 4 \log \frac{4}{3} \Delta \mathcal{E}^{\varepsilon, a, A},$$

analogously to Lemma 4.2.1. This can only be satisfied if the number  $|J|$  of elements in  $J$  is less than  $K$ . Thus we can pick at least  $K$  cross-sections  $s_k \in J_k$  with the desired property.  $\square$

**Proposition 4.2.4.** *In  $n = 3$  dimensions, for  $\varepsilon < 1$  and  $a > 1$  we have*

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A) \min \left\{ \sqrt{\varepsilon} \left( \sqrt{a} + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| \right), a-1 \right\}.$$

*Proof.* Let us assume

$$\Delta \mathcal{E}^{\varepsilon, a, A} \leq 2^{-13} \mathcal{H}^2(A) \min\{1, a-1\}.$$

We now consider cross-sections  $\{x_3 = s_k\}$ ,  $k = 1, \dots, K$ , where  $\frac{1}{2} \geq s_1 > \dots > s_K \geq 2^{-2K-1}$  denote the  $K$  selected cross-sections from the previous lemma and  $K \geq 1$  is chosen such that

$$128 \frac{a}{a-1} \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\mathcal{H}^2(A)} \in [2^{-2K-2}, 2^{-2K-1}],$$

which is possible since  $\Delta \mathcal{E}^{\varepsilon, a, A} \leq 2^{-13} \mathcal{H}^2(A) \min\{a-1, 1\}$ . The total amount of particles bypassing  $\Sigma^*$  on at least one of the cross-sections  $\{x_3 = s_k\}$ ,  $k = 1, \dots, K$ , can be bounded above by

$$\sum_{k=1}^K \Delta F_{s_k} \leq \frac{4 \Delta \mathcal{E}^{\varepsilon, a, A}}{K(a-1)} \sum_{k=1}^K \frac{1}{s_k} \leq \frac{16 \Delta \mathcal{E}^{\varepsilon, a, A}}{K(a-1)s_K} \leq \frac{1}{8} \mathcal{H}^2(A).$$

Likewise, we can bound the set  $\Lambda_k \subset \Gamma$  of particles that between the cross-sections  $\{x_3 = s_{k+1}\}$  and  $\{x_3 = s_k\}$  travel horizontally by more than  $s_k - s_{k+1}$ . Indeed, we have

$$\begin{aligned} \mathcal{E}^{\varepsilon, a, A}[\chi^*] &= \int_{\Gamma \times I} r_{\chi^*}^{\varepsilon, a}(\chi^*(p, t)) |\dot{\chi}^*(p, t)| \, dt \, dP_{\Gamma}(p) \geq \int_{\Gamma} \mathcal{H}^1(\chi_p^*(I)) \, dP_{\Gamma}(p) \\ &\geq \int_{\Gamma} 1 + (\sqrt{2} - 1)(s_k - s_{k+1}) \mathbf{1}_{\Lambda_k}(p) \, dP_{\Gamma}(p) = P_{\Gamma}(\Gamma) + (\sqrt{2} - 1)(s_k - s_{k+1}) P_{\Gamma}(\Lambda_k) \end{aligned}$$

so that  $\Delta \mathcal{E}^{\varepsilon, a, A} = \mathcal{E}^{\varepsilon, a, A}[\chi^*] - P_{\Gamma}(\Gamma) \geq (\sqrt{2} - 1)(s_k - s_{k+1}) P_{\Gamma}(\Lambda_k)$ . Thus,

$$\sum_{k=1}^K P_{\Gamma}(\Lambda_k) \leq \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\sqrt{2} - 1} \sum_{k=1}^K \frac{1}{s_k - s_{k+1}} \leq \frac{8 \Delta \mathcal{E}^{\varepsilon, a, A}}{(\sqrt{2} - 1)s_K} \leq \frac{\mathcal{H}^2(A)}{4}.$$

In the following, we will only consider particles that on each cross-section flow through  $\Sigma^*$  and that between any  $\{x_3 = s_{k+1}\}$  and  $\{x_3 = s_k\}$  travel horizontally by no more than  $s_k - s_{k+1}$ , that is, particles in

$$\bar{\Gamma} = \bigcap_{k=1}^K (\Gamma_{s_k} \setminus \Lambda_k) \quad \text{with} \quad P_{\Gamma}(\bar{\Gamma}) \geq P_{\Gamma}(\Gamma) - \sum_{k=1}^K \Delta F_{s_k} - \sum_{k=1}^K P_{\Gamma}(\Lambda_k) \geq \frac{5}{8} \mathcal{H}^2(A).$$

Now, for  $s \in [0, 1]$  let

$$\tilde{\mu}_s = \xi^*(\cdot, s)_{\#} (P_{\Gamma} \llcorner \bar{\Gamma})$$

be the mass flux distribution on  $\{x_3 = s\}$  of all above selected particles. We have  $\|\tilde{\mu}_0\|_{\text{fbm}} = \|\tilde{\mu}_1\|_{\text{fbm}} = \|\tilde{\mu}_{s_1}\|_{\text{fbm}} = \dots = \|\tilde{\mu}_{s_K}\|_{\text{fbm}} \geq \frac{5}{8} \mathcal{H}^2(A)$ . The total urban planning cost can now be estimated as follows (see Remark 3.1.4),

$$\begin{aligned} \mathcal{E}^{\varepsilon, a, A}[\chi^*] &\geq \inf_{\chi} \mathcal{E}^{\varepsilon, a, \mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1}[\chi] + \inf_{\chi} \mathcal{E}^{\varepsilon, a, \tilde{\mu}_0, \tilde{\mu}_{s_K}}[\chi] + \sum_{k=1}^{K-1} \inf_{\chi} \mathcal{E}^{\varepsilon, a, \tilde{\mu}_{s_{k+1}}, \tilde{\mu}_{s_k}}[\chi] + \inf_{\chi} \mathcal{E}^{\varepsilon, a, \tilde{\mu}_{s_1}, \tilde{\mu}_1}[\chi] \\ &\geq W_1(\mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1) + W_1(\tilde{\mu}_0, \tilde{\mu}_{s_K}) + \sum_{k=1}^{K-1} W_1(\tilde{\mu}_{s_{k+1}}, \tilde{\mu}_{s_k}) + W_1(\tilde{\mu}_{s_1}, \tilde{\mu}_1) \end{aligned}$$

so that the excess energy is bounded below by

$$\begin{aligned}
\Delta \mathcal{E}^{\varepsilon, a, A} &= \mathcal{E}^{\varepsilon, a, A}[\chi^*] - \mathcal{H}^2(A) \\
&= \mathcal{E}^{\varepsilon, a, A}[\chi^*] - \|\mu_0 - \tilde{\mu}_0\|_{\text{fbm}} - \|\tilde{\mu}_0\|_{\text{fbm}} \left[ s_K + \sum_{k=1}^{K-1} (s_k - s_{k+1}) + 1 - s_1 \right] \\
&\geq W_1(\tilde{\mu}_{s_K}, \tilde{\mu}_0) - \|\tilde{\mu}_{s_K}\| s_K + \sum_{k=1}^{K-1} (W_1(\tilde{\mu}_{s_k}, \tilde{\mu}_{s_{k+1}}) - \|\tilde{\mu}_{s_k}\| (s_k - s_{k+1})) ,
\end{aligned}$$

where in the last step we used  $W_1(\mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1) \geq \|\mu_0 - \tilde{\mu}_0\|_{\text{fbm}}$  as well as  $W_1(\tilde{\mu}_{s_1}, \tilde{\mu}_1) \geq \|\tilde{\mu}_0\|_{\text{fbm}}(1 - s_1)$ . Now let  $(\cdot)'$  denote the projection onto  $\mathbb{R}^2$ , that is,

$$(\cdot)' : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad x' = (x_1, x_2) \quad \text{and} \quad \mu' = (\cdot)'_{\#} \mu$$

for  $\mu \in \text{fbm}(\mathbb{R}^3)$ . Let us define also

$$\tilde{\Pi}(\tilde{\mu}_s, \tilde{\mu}_t) = \left\{ \mu \in \text{fbm}(\mathbb{R}^3 \times \mathbb{R}^3) : \begin{array}{l} \mu(A \times \mathbb{R}^3) = \tilde{\mu}_s(A), \mu(\mathbb{R}^3 \times B) = \tilde{\mu}_t(B) \text{ for all } A, B \in \mathcal{B}(\mathbb{R}^3) \\ \text{spt} \mu \subseteq \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 : |x' - y'| \leq |x_3 - y_3|\} \end{array} \right\}.$$

Using the inequality

$$\sqrt{\alpha^2 + \beta^2} - \beta \geq \frac{\alpha^2}{4\beta} \quad \text{for all } \frac{\alpha}{\beta} \in [0, \sqrt{8}]$$

and the notation from Definition 1.1.1 we get

$$\begin{aligned}
W_1(\tilde{\mu}_s, \tilde{\mu}_t) - \|\tilde{\mu}_s\| (s - t) &= \inf_{\mu \in \tilde{\Pi}(\tilde{\mu}_s, \tilde{\mu}_t)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} |x - y| - (s - t) \, d\mu(x, y) \\
&= \inf_{\mu \in \tilde{\Pi}(\tilde{\mu}_s, \tilde{\mu}_t)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} \sqrt{|x' - y'|^2 + (s - t)^2} - (s - t) \, d\mu(x, y) \\
&\geq \inf_{\mu \in \tilde{\Pi}(\tilde{\mu}_s, \tilde{\mu}_t)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|x' - y'|^2}{4(s - t)} \, d\mu(x, y) \\
&\geq \frac{W_2(\tilde{\mu}'_s, \tilde{\mu}'_t)^2}{4(s - t)}
\end{aligned}$$

for any  $s, t \in [0, 1]$ . Thus, the previous estimate becomes

$$\begin{aligned}
\Delta \mathcal{E}^{\varepsilon, a, A} &\geq \sum_{k=1}^{K-1} \frac{W_2(\tilde{\mu}'_{s_k}, \tilde{\mu}'_{s_{k+1}})^2}{4(s_k - s_{k+1})} + \frac{W_2(\tilde{\mu}'_{s_K}, \tilde{\mu}'_0)^2}{4s_K} \\
&\geq \sum_{k=1}^{K-1} \frac{\left[ W_2(\tilde{\mu}'_{s_k}, \tilde{\mu}'_0) - W_2(\tilde{\mu}'_0, \tilde{\mu}'_{s_{k+1}}) \right]^2}{4(s_k - s_{k+1})} + \frac{W_2(\tilde{\mu}'_{s_K}, \tilde{\mu}'_0)^2}{4s_K},
\end{aligned}$$

where we have used the triangle inequality. Now Jensen's inequality, Lemma 4.1.1, and Lemma 4.2.3 imply

$$W_2(\tilde{\mu}'_{s_k}, \tilde{\mu}'_0)^2 \geq \left( \frac{W_1(\tilde{\mu}'_{s_k}, \tilde{\mu}'_0)}{\sqrt{\|\tilde{\mu}'_{s_k}\|_{\text{fbm}}}} \right)^2 \geq C^2 \frac{\|\tilde{\mu}'_{s_k}\|_{\text{fbm}}^2}{\omega_2 N_{s_k}} \geq \frac{25C^2 K \mathcal{H}^2(A)^2 \varepsilon_{s_k}}{256 \omega_2 \Delta \mathcal{E}^{\varepsilon, a, A}} =: (d_k^*)^2,$$

thus the bound on  $\Delta \mathcal{E}^{\varepsilon, a, A}$  can be expressed as

$$\Delta \mathcal{E}^{\varepsilon, a, A} \geq \inf_{d_k \geq d_k^*, k=1, \dots, K} \sum_{k=1}^{K-1} \frac{[d_k - d_{k+1}]^2}{4(s_k - s_{k+1})} + \frac{d_K^2}{4s_K}.$$

Using then first order optimality conditions one can check that this convex optimisation problem is minimised by the choice  $d_k = d_k^*$ ,  $k = 1, \dots, K$ , so that

$$\begin{aligned} \Delta \mathcal{E}^{\varepsilon, a, A} &\geq \sum_{k=1}^{K-1} \frac{[d_k^* - d_{k+1}^*]^2}{4(s_k - s_{k+1})} + \frac{(d_K^*)^2}{4s_K} \sim \frac{K \mathcal{H}^2(A)^2 \varepsilon}{\Delta \mathcal{E}^{\varepsilon, a, A}} \left[ \sum_{k=1}^{K-1} \frac{(\sqrt{s_k} - \sqrt{s_{k+1}})^2}{s_k - s_{k+1}} + 1 \right] \\ &= \frac{K \mathcal{H}^2(A)^2 \varepsilon}{\Delta \mathcal{E}^{\varepsilon, a, A}} \left[ \sum_{k=1}^{K-1} \frac{1 - \sqrt{\frac{s_{k+1}}{s_k}}}{1 + \sqrt{\frac{s_{k+1}}{s_k}}} + 1 \right] \gtrsim \frac{\mathcal{H}^2(A)^2 \varepsilon}{\Delta \mathcal{E}^{\varepsilon, a, A}} \left| \log \left( 128 \frac{a}{a-1} \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\mathcal{H}^2(A)} \right) \right|^2 \end{aligned}$$

using  $2^{-K} \geq 128 \frac{a}{a-1} \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\mathcal{H}^2(A)}$ . Abbreviating  $z = 128 \frac{a}{a-1} \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\mathcal{H}^2(A)}$ , this inequality can be transformed into  $\frac{z^2}{|\log z|^2} \gtrsim \left( \frac{a}{a-1} \right)^2 \varepsilon$ , which can be solved for  $z$  to yield  $z \gtrsim \frac{a}{a-1} \sqrt{\varepsilon} \max \left\{ 1, \left| \log \frac{a\sqrt{\varepsilon}}{a-1} \right| \right\}$ . Thus,

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A) \sqrt{\varepsilon} \max \left\{ 1, \left| \log \frac{a\sqrt{\varepsilon}}{a-1} \right| \right\}.$$

This bound was derived under the assumption that  $\Delta \mathcal{E}^{\varepsilon, a, A} \leq 2^{-13} \mathcal{H}^2(A) \min\{1, a-1\}$  so that together we obtain

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A) \min \left\{ \sqrt{\varepsilon} \left( 1 + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| \right), a-1, 1 \right\}.$$

In addition we will derive the bound

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^{n-1}(A) \min \left\{ \sqrt{a^2 - 1}^{\frac{n-2}{n-1}} \varepsilon^{\frac{1}{n-1}}, a-1 \right\} = \mathcal{H}^2(A) \min \left\{ \sqrt{\varepsilon}^4 \sqrt{a^2 - 1}, a-1 \right\}$$

in Section 4.2.4. Now assume, the minimum in the former bound were achieved by 1, that is,  $a-1 > 1$  as well as  $\sqrt{\varepsilon} \left( 1 + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| \right) > 1$ , and the estimate reads  $\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A)$ . As long as  $\sqrt{\varepsilon}^4 \sqrt{a^2 - 1} < 1$ , the latter bound does not overrule this estimate. However, this can only happen if  $\varepsilon \sim 1$  and  $a-1 \sim 1$  so that actually  $\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A)(a-1)$ . Thus, the 1 in the former bound may be neglected, and we finally arrive at

$$\begin{aligned} \Delta \mathcal{E}^{\varepsilon, a, A} &\gtrsim \mathcal{H}^2(A) \max \left\{ \min \left\{ \sqrt{\varepsilon} \left( 1 + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| \right), a-1 \right\}, \min \left\{ \sqrt{\varepsilon}^4 \sqrt{a^2 - 1}, a-1 \right\} \right\} \\ &\gtrsim \mathcal{H}^2(A) \min \left\{ \sqrt{\varepsilon} \left( 1 + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| + \sqrt[4]{a^2 - 1} \right), a-1 \right\} \\ &\gtrsim \mathcal{H}^2(A) \min \left\{ \sqrt{\varepsilon} \left( \sqrt{a} + \left| \log \frac{a-1}{\sqrt{\varepsilon}} \right| \right), a-1 \right\}. \end{aligned}$$

□

### 4.2.3 Lower bound in $nD$ for $n > 3$

In two dimensions we bounded the excess energy from below using the relaxed energy between  $\mu_0, \mu_1$ , and a generic cross-section  $\{x_2 = s\}$  with  $s \sim \frac{1}{2}$ . In three dimensions we employed the same technique, only now considering various cross-sections. In higher dimensions, where the energy will be concentrated near the boundary, we apply the technique to a cross-section near the boundary. The next lemma is a generalisation of Lemma 4.2.1.

**Lemma 4.2.5.** *Let  $T \in (0, 1)$ . There is a generic  $s \in (0, T)$  such that*

$$N_s \leq \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\varepsilon T} \quad \text{and} \quad \Delta F_s \leq \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{(a-1)T}.$$

*Proof.* If almost all  $s \in (0, T)$  satisfied  $N_s > \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{\varepsilon T}$  or  $\Delta F_s > \frac{\Delta \mathcal{E}^{\varepsilon, a, A}}{(a-1)T}$ , then we had  $\Delta \mathcal{E}^{\varepsilon, a, A} \geq \int_0^T \varepsilon N_s + (a-1) \Delta F_s \, dt > \Delta \mathcal{E}^{\varepsilon, a, A}$ , a contradiction. □

**Proposition 4.2.6.** *For  $\varepsilon < 1$  and  $a > 1$  we have*

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^2(A) \min \left\{ \sqrt{a} \sqrt{a-1}^{\frac{n-3}{n-1}} \varepsilon^{\frac{1}{n-1}}, a-1 \right\}.$$

*Proof.* The lower bound proof now is as follows (note how the procedure parallels that of the case  $n = 2$ ). Assume  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq \frac{1}{4}\mathcal{H}^{n-1}(A)(a-1)$  and pick

$$T = \frac{2\Delta\mathcal{E}^{\varepsilon,a,A}}{\mathcal{H}^{n-1}(A)(a-1)}.$$

Let  $s \in (0, T)$  be the cross-section from the previous lemma. Let  $\tilde{\mu}_0, \tilde{\mu}_1$  be the irrigating and irrigated measure of all particles in  $\Gamma_s$  according to (4.2.7) and (4.2.8). Now

$$\begin{aligned} \mathcal{E}^{\varepsilon,a,A}[\chi^*] &\geq \inf_{\chi} \mathcal{E}^{\varepsilon,a,\tilde{\mu}_0,\mu_s}[\chi] + \inf_{\chi} \mathcal{E}^{\varepsilon,a,\mu_s,\tilde{\mu}_1}[\chi] + W_1(\mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1) \\ &\geq W_1(\tilde{\mu}_0, \mu_s) + W_1(\mu_s, \tilde{\mu}_1) + W_1(\mu_0 - \tilde{\mu}_0, \mu_1 - \tilde{\mu}_1) \\ &\geq s\|\mu_s\|_{\text{fbm}} + C\|\mu_s\|_{\text{fbm}} \min\left\{\frac{R^2}{s}, R\right\} + (1-s)\|\mu_s\|_{\text{fbm}} + \|\mu_0 - \tilde{\mu}_0\|_{\text{fbm}} \\ &\geq \mathcal{E}^{*,a,A} + CR\|\mu_s\|_{\text{fbm}} \min\left\{\frac{R}{T}, 1\right\}, \end{aligned}$$

where we have employed Lemma 4.1.1 with  $R = \sqrt[n-1]{\frac{\|\mu_s\|_{\text{fbm}}}{N_s\omega_{n-1}}}$ . By virtue of the previous lemma, our assumption on  $\Delta\mathcal{E}^{\varepsilon,a,A}$ , and our choice of  $T$ ,  $\|\mu_s\|_{\text{fbm}} = \mathcal{H}^{n-1}(A) - \Delta F_s \geq \frac{\mathcal{H}^{n-1}(A)}{2}$  and  $N_s \leq \frac{\mathcal{H}^{n-1}(A)(a-1)}{2\varepsilon}$  so that  $R \geq \sqrt[n-1]{\frac{\varepsilon}{\omega_{n-1}(a-1)}}$ . Inserting everything into the above estimate,

$$\Delta\mathcal{E}^{\varepsilon,a,A} \geq C \frac{\mathcal{H}^{n-1}(A)}{2} \sqrt[n-1]{\frac{\varepsilon}{\omega_{n-1}(a-1)}} \min\left\{\frac{\mathcal{H}^{n-1}(A)}{2 \sqrt[n-1]{\omega_{n-1}} \Delta\mathcal{E}^{\varepsilon,a,A}} \varepsilon^{\frac{1}{n-1}} (a-1)^{\frac{n-2}{n-1}}, 1\right\}. \quad (4.2.9)$$

Now we consider two cases.

- If  $\Delta\mathcal{E}^{\varepsilon,a,A} \leq \frac{\mathcal{H}^{n-1}(A)}{2 \sqrt[n-1]{\omega_{n-1}}} \varepsilon^{\frac{1}{n-1}} (a-1)^{\frac{n-2}{n-1}}$ , inequality (4.2.9) yields

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A) \sqrt[n-1]{\frac{\varepsilon}{a-1}}.$$

Note that this is only possible for  $a-1 \gtrsim 1$ .

- If  $\Delta\mathcal{E}^{\varepsilon,a,A} > \frac{\mathcal{H}^{n-1}(A)}{2 \sqrt[n-1]{\omega_{n-1}}} \varepsilon^{\frac{1}{n-1}} (a-1)^{\frac{n-2}{n-1}}$ , inequality (4.2.9) can be solved for  $\Delta\mathcal{E}^{\varepsilon,a,A}$  to give

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A) \varepsilon^{\frac{1}{n-1}} \sqrt{a-1}^{\frac{n-3}{n-1}}.$$

Both cases can be summarised as

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A) \varepsilon^{\frac{1}{n-1}} \cdot \begin{cases} (a-1)^{-\frac{1}{n-1}} & \text{if } a-1 > 1, \\ \sqrt{a-1}^{\frac{n-3}{n-1}} & \text{else.} \end{cases}$$

This bound was derived under the assumption that  $\Delta\mathcal{E}^{\varepsilon,a,A} \lesssim \mathcal{H}^{n-1}(A)(a-1)$ . Together with the bound  $\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A) \min\{\sqrt{a^2-1}^{\frac{n-2}{n-1}} \varepsilon^{\frac{1}{n-1}}, a-1\}$  derived in Section 4.2.4, the estimates can be combined into

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^2(A) \min\left\{\sqrt{a}\sqrt{a-1}^{\frac{n-3}{n-1}} \varepsilon^{\frac{1}{n-1}}, a-1\right\},$$

as can be easily checked for the different cases  $a-1 \ll 1$ ,  $a-1 \sim 1$ , and  $a-1 \gg 1$ .  $\square$

#### 4.2.4 Boundary contribution

Here we estimate the energy associated with the particles travelling from the boundary to the network  $\Sigma^*$ . The estimate holds for all  $n \geq 3$ . We need a few preparatory lemmas, the first of which is a refinement of [BOS02, Thm. 3.16].

**Lemma 4.2.7.** *Let  $S \subset \mathbb{R}^n$  be connected with  $\mathcal{H}^1(S) = l$  and  $K \subset \mathbb{R}^n$  with  $\mathcal{H}^n(K) = V$ . Then there is a positive constant  $C \equiv C(n)$  such that*

$$r_{S,K} := \frac{1}{V} \int_K \text{dist}(x, S) d\mathcal{H}^n(x) \geq C \min\left\{V^{\frac{1}{n}}, \left(\frac{V}{l}\right)^{\frac{1}{n-1}}\right\},$$

where in the case  $V = 0$  the normalised integral is to be interpreted as  $r_{S,K} = 0$ .

*Proof.* Let us pick  $\delta = \zeta \left(\frac{V}{l}\right)^{\frac{1}{n-1}}$  for some constant  $\zeta \equiv \zeta(n) \in (0, 1)$  to be specified later. Now we cover  $K$  and the surrounding space by a grid of cubes with side length  $\delta$ . In [BOS02, Lem. 3.17] the authors prove that  $S$  intersects at most  $k = (2^n + 1)\frac{l}{\delta} + (2^n + 1)$  cubes by noting that one needs a curve segment of at least length  $\delta$  to intersect any union of  $2^n + 1$  cubes. Those  $k$  cubes together have a volume of

$$V_S = k\delta^n = (2^n + 1)\zeta^{n-1}V + (2^n + 1)\zeta^n \left(\frac{V}{l}\right)^{\frac{n}{n-1}}.$$

We shall call a cube a neighbour of  $S$  if it is a neighbour of a cube intersecting  $S$ , and we shall call a cube a second neighbour if it is adjacent to a neighbour of  $S$ . We distinguish three cases:

- If  $(2^n + 1)\zeta^{n-1}V \leq (2^n + 1)\zeta^n \left(\frac{V}{l}\right)^{\frac{n}{n-1}}$  or equivalently  $l \leq \zeta^{1-\frac{1}{n}}V^{\frac{1}{n}}$ , then the average distance  $r_{S,K}$  of points in  $K$  to  $S$  is greater than  $V^{\frac{1}{n}}$  up to a constant factor. Indeed, it is clear that for  $\xi \geq 0$  we have  $r_{\xi S, \xi K} = \xi r_{S,K}$ . Furthermore,

$$r_1 := \inf_{\substack{\tilde{K} \subset \mathbb{R}^n, \mathcal{H}^n(\tilde{K})=1 \\ \tilde{S} \subset \mathbb{R}^n \text{ connected, } \mathcal{H}^1(\tilde{S}) \leq \zeta^{1-\frac{1}{n}}}} \int_{\tilde{K}} \text{dist}(x, \tilde{S}) d\mathcal{H}^n(x) > 0 \quad (4.2.10)$$

so that  $r_{S,K} = V^{\frac{1}{n}} r_{V^{-\frac{1}{n}}S, V^{-\frac{1}{n}}K} \geq V^{\frac{1}{n}} r_1$ . The positivity of  $r_1$  will be proved in Lemma 4.2.9 below.

- If  $l > \zeta^{1-\frac{1}{n}}V^{\frac{1}{n}}$  and if the intersected cubes, the neighbours, and the second neighbours contain less than  $\frac{V}{2}$  of  $K$ , then at least  $\frac{V}{2}$  of  $K$  is at distance larger than  $2\delta$  from  $S$  so that  $r_{S,K} \geq (\frac{V}{2}(2\delta) + \frac{V}{2}0)/V = \delta$ .
- Now let  $l > \zeta^{1-\frac{1}{n}}V^{\frac{1}{n}}$  and let the intersected cubes and first and second neighbours contain more than  $\frac{V}{2}$  of  $K$ . The total number of intersecting cubes and neighbours is less than  $k3^n$ , where  $3^n$  is the number of neighbours a cube has (including the cube itself). Thus, the second neighbours actually intersect  $K$  on a volume of more than

$$\frac{V}{2} - k3^n\delta^n = \frac{V}{2} - 3^n V_S \geq \frac{V}{2} - 3^n \cdot 2(2^n + 1)\zeta^{n-1}V \geq \frac{V}{4}$$

if we choose  $\zeta$  small enough. Thus,  $r_{S,K} \geq \frac{\delta}{4}$ . □

*Remark 4.2.8.* Intuitively it is clear how in the previous lemma  $S$  and  $K$  have to be arranged to give the smallest  $r_{S,K}$ ;  $S$  should be chosen as a straight line of length  $l$  and  $K = B_r(S)$ , where  $r$  is such that  $\mathcal{H}^n(K) = V$ . If  $S$  is very short, then  $K$  is almost a ball, and the average distance to  $S$  behaves like the average distance to the ball centre, which scales like  $V^{\frac{1}{n}}$ . If  $S$  is very long, then  $K$  is almost a cylinder, and the average distance to its midaxis scales like  $\left(\frac{V}{l}\right)^{\frac{1}{n-1}}$ .

**Lemma 4.2.9.** *Let  $r_1$  be given by equation (4.2.10), then  $r_1 > 0$ .*

*Proof.* For a contradiction assume there exist sequences  $\tilde{K}_i, \tilde{S}_i$  with  $\mathcal{H}^n(\tilde{K}_i) = 1, \mathcal{H}^1(\tilde{S}_i) \leq \zeta^{1-\frac{1}{n}}$  such that

$$F(\tilde{K}_i, \tilde{S}_i) = \int_{\tilde{K}_i} \text{dist}(x, \tilde{S}_i) d\mathcal{H}^n(x) \rightarrow 0$$

as  $i \rightarrow \infty$ , where without loss of generality we may assume the  $\tilde{K}_i, \tilde{S}_i$  to be uniformly bounded. Thanks to Golab's Theorem we may assume that, up to a subsequence,  $\tilde{S}_i \rightarrow S$  in the Hausdorff sense and that  $\mathcal{H}^1(S) \leq \zeta^{1-\frac{1}{n}}$ . Moreover,

$$\inf_{\tilde{K}, \mathcal{H}^n(\tilde{K})=1} F(\tilde{K}, S) = 0, \quad (4.2.11)$$

since

$$\begin{aligned} F(\tilde{K}_i, S) &= \int_{\tilde{K}_i} \text{dist}(x, S) d\mathcal{H}^n(x) \geq \int_{\tilde{K}_i} \text{dist}(x, \tilde{S}_i) - \text{dist}(S, \tilde{S}_i) d\mathcal{H}^n(x) \\ &= F(\tilde{K}_i, \tilde{S}_i) - \text{dist}(S, \tilde{S}_i) \mathcal{H}^n(\tilde{K}_i) \rightarrow 0. \end{aligned}$$

This leads to a contradiction. In fact, let  $K \subseteq \mathbb{R}^n$  with  $\mathcal{H}^n(K) = 1$  and set  $K_\delta = \{x \in K : \text{dist}(x, S) > \delta\}$  as well as  $S_\delta = \{x \in \mathbb{R}^n : \text{dist}(x, S) \leq \delta\}$ . Then,

$$F(K, S) \geq \delta \mathcal{H}^n(K_\delta) = \delta(1 - \mathcal{H}^n(K \setminus K_\delta)) \geq \delta(1 - \mathcal{H}^1(S_\delta)).$$

Now [BS14, Lemma 4.1] implies  $\mathcal{H}^n(S_\delta) \leq C \mathcal{H}^1(S) \delta^{n-1}$  for some constant  $C \equiv C(n)$  so that we can choose  $\delta > 0$  such that  $\mathcal{H}^n(S_\delta) < \frac{1}{2}$ . Then  $F(K, S) \geq \delta(1 - \frac{1}{2}) > 0$  independent of  $K$ , which contradicts equation (4.2.11).  $\square$

The next lemma is a regularity result for the optimal irrigation pattern  $\chi^*$  and corresponding network  $\Sigma^*$ .

**Lemma 4.2.10.** *There are an optimal irrigation pattern  $\chi^*$  and corresponding network  $\Sigma^*$  such that  $\Sigma^*$  is symmetric across  $\{x_n = \frac{1}{2}\}$  and  $\chi_p^*(I) \cap \Sigma^*$  is connected for almost all  $p \in \Gamma$ .*

*Proof.* We may assume  $\chi^*$  to have the single path property (see [BW15, Proposition 3.4.1]). Furthermore, due to the problem symmetry we may assume  $\chi_p^*(I)$  to be symmetric across  $\{x_n = \frac{1}{2}\}$  for all  $p \in \Gamma$ . Indeed, suppose without loss of generality and potentially reparameterising the pattern that  $\chi_p^*(\frac{1}{2}) \in \{x_n = \frac{1}{2}\}$  for all  $p \in \Gamma$ . We now consider a pattern with the property that the position of a particle  $p$  at time  $\frac{1}{2} + t$  is the same as at time  $\frac{1}{2} - t$ , but reflected across  $\{x_n = \frac{1}{2}\}$ . As an explicit formula,

$$\chi(p, t) = \begin{cases} \chi^*(p, t) & t \in [0, \frac{1}{2}], \\ ((\chi^*(p, 1-t))_1, \dots, (\chi^*(p, 1-t))_{n-1}, 1 - (\chi^*(p, 1-t))_n) & t \in [\frac{1}{2}, 1]. \end{cases}$$

If  $\chi^*$  is optimal, then this new symmetric pattern must also be optimal with the same cost. The corresponding  $\Sigma^*$  is also symmetric.

Now assume  $\chi_p^*(I) \cap \Sigma^*$  to be not connected. Then there are  $s_1, s_2, \delta > 0$  with  $\chi_p^*(t) \notin \Sigma^*$  for  $t \in (s_1, s_2)$  and  $\chi_p^*(t) \in \Sigma^*$  for  $t \in (s_1 - \delta, s_1) \cup (s_2, s_2 + \delta)$ . Let us abbreviate  $\hat{\Sigma} = \chi_p^*((s_1 - \delta, s_1)) \subset \Sigma^*$ . Without loss of generality, assume  $\hat{\Sigma} \subset \{x_n < \frac{1}{2}\}$  and  $\mathcal{H}^1(\hat{\Sigma}) > 0$ .

Let  $[x]_{\chi^*} \subset \Gamma$  with  $P_\Gamma([x]_{\chi^*}) = m_{\chi^*}(x)$  denote the set of particles flowing through  $x \in \hat{\Sigma}$ . Note  $m_{\chi^*}(x) > \frac{\varepsilon}{a-1}$  by the relation between  $\chi^*$  and  $\Sigma^*$  from Section 2.2. Due to the single path property, all particles  $q \in [x]_{\chi^*}$  follow the same path between  $x$  and its reflection  $\bar{x} = \chi_p^*(s_3)$  across  $\{x_n = \frac{1}{2}\}$ , in other words,  $\chi_p^*((s_1, s_3)) \subset \chi_q^*(I)$  for all  $q \in [x]_{\chi^*}$ . Thus,  $m_{\chi^*}(x) > \frac{\varepsilon}{a-1}$  for all  $x \in \chi_p^*((s_1, s_3))$  so that  $\chi_p^*((s_1, s_3)) \subset \Sigma^*$ . This however contradicts  $\chi_p^*((s_1, s_2)) \subset \mathbb{R}^n \setminus \Sigma^*$ .  $\square$

**Proposition 4.2.11** (Boundary contribution). *In  $n \geq 3$  dimensions the contribution to the excess energy of particles travelling from the initial distribution  $\mu_0$  to the optimal network  $\Sigma^*$  can be estimated as*

$$\Delta \mathcal{E}^{\varepsilon, a, A} \gtrsim \mathcal{H}^{n-1}(A) \min \left\{ a - 1, \sqrt{a^2 - 1}^{\frac{n-2}{n-1}} \varepsilon^{\frac{1}{n-1}} \right\}.$$

*Proof.* By Lemma 4.2.10 we may assume that the path  $\chi_p^*(I)$  of any particle  $p \in \Gamma$  can be subdivided into three connected segments, one outside  $\Sigma^*$ , one within, and again one outside. Let

$$\begin{aligned} t(p) &= \inf\{t \in I : \chi_p^*(t) \in \Sigma^*\}, \\ x(p) &= \chi_p^*(t(p)), \\ r(p) &= |(x(p) - \chi_p^*(0))'| \end{aligned}$$

denote the time and position at which particle  $p \in \Gamma$  reaches  $\Sigma^*$  for the first time, as well as the accumulated horizontal motion until then. We have

$$\begin{aligned} \int_I r_{\chi^*}^{\varepsilon, a}(\chi^*(p, t)) |\dot{\chi}^*(p, t)| dt &\geq \int_0^{t(p)} a |\dot{\chi}^*(p, t)| dt + \int_{t(p)}^1 |\dot{\chi}^*(p, t)| dt \\ &\geq a \sqrt{(r(p))^2 + ((x(p))_n)^2} + (1 - (x(p))_n) \geq 1 + \min\{\sqrt{a^2 - 1} r(p), a - 1\} \end{aligned}$$



upon optimising for  $(x(p))_n \in [0, 1]$ , which would yield  $(x(p))_n = \min\{\frac{r(p)}{\sqrt{a^2-1}}, 1\}$ . Let  $\tilde{\Gamma} \subset \Gamma$  denote the particles for which  $r(p) \leq \sqrt{\frac{a-1}{a+1}}$ . The excess energy can now be estimated below by

$$\begin{aligned}\Delta\mathcal{E}^{\varepsilon,a,A} &= \int_{\Gamma} \int_I r_{\chi^*}^{\varepsilon,a}(\chi^*(p,t)) |\dot{\chi}^*(p,t)| dt dP_{\Gamma}(p) - P_{\Gamma}(\Gamma) \\ &\geq \int_{\tilde{\Gamma}} 1 + \sqrt{a^2 - 1} r(p) dP_{\Gamma}(p) + \int_{\Gamma \setminus \tilde{\Gamma}} 1 + a - 1 dP_{\Gamma}(p) - P_{\Gamma}(\Gamma) \\ &= \sqrt{a^2 - 1} P_{\Gamma}(\tilde{\Gamma}) + (a - 1) P_{\Gamma}(\Gamma \setminus \tilde{\Gamma}),\end{aligned}$$

where  $r = \frac{1}{P_{\Gamma}(\tilde{\Gamma})} \int_{\tilde{\Gamma}} r(p) dP_{\Gamma}(p)$  denotes the average horizontal motion of particles in  $\tilde{\Gamma}$  before they reach  $\Sigma^*$ . Now combining Lemmas 4.2.1 and 4.2.10 we know that  $\Sigma^*$  intersects  $\{x_n = \frac{1}{2}\}$  in at most  $N = \frac{\Delta\mathcal{E}^{\varepsilon,a,A}}{\varepsilon}$  points  $x_1, \dots, x_N$ . All particles in  $\tilde{\Gamma}$  flow through such a point, and  $\tilde{\Gamma} \subset \bigcup_{i=1}^N [x_i]_{\chi^*}$  using the notation from Definition 2.1.5. Denote by  $\Sigma_i \subset \Sigma^*$  the connected component of  $\Sigma^*$  containing  $x_i$  (by the single path property of the minimiser from Lemma 4.2.10, the  $\Sigma_i$  are disjoint). If  $\Sigma'_i$  denotes the orthogonal projection of  $\Sigma_i$  onto  $\{x_n = 0\}$ , we have

$$r \geq \frac{1}{P_{\Gamma}(\tilde{\Gamma})} \sum_{i=1}^N \int_{[x_i]_{\chi^*} \cap \tilde{\Gamma}} \text{dist}(\chi_p^*(0), \Sigma'_i) dP_{\Gamma}(p).$$

Using Lemma 4.2.7 in dimension  $n - 1$  and for  $S = \Sigma'_i$  and  $K = \{\chi_p^*(0) : p \in [x_i]_{\chi^*} \cap \tilde{\Gamma}\}$  we thus obtain

$$\begin{aligned}r &\geq \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \sum_{i=1}^N P_{\Gamma}([x_i]_{\chi^*} \cap \tilde{\Gamma}) \min \left\{ P_{\Gamma}([x_i]_{\chi^*} \cap \tilde{\Gamma})^{\frac{1}{n-1}}, \left( \frac{P_{\Gamma}([x_i]_{\chi^*} \cap \tilde{\Gamma})}{\mathcal{H}^1(\Sigma_i)} \right)^{\frac{1}{n-2}} \right\} \\ &\geq \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \inf_{\substack{V_1, \dots, V_N \geq 0, \\ l_1, \dots, l_N \geq 0, \\ V_1 + \dots + V_N = P_{\Gamma}(\tilde{\Gamma}), \\ l_1 + \dots + l_N \leq \mathcal{H}^1(\Sigma^*)}} \sum_{i=1}^N V_i \min \left\{ V_i^{\frac{1}{n-1}}, \left( \frac{V_i}{l_i} \right)^{\frac{1}{n-2}} \right\}.\end{aligned}$$

Without loss of generality assume  $V_i^{\frac{1}{n-1}} \leq \left( \frac{V_i}{l_i} \right)^{\frac{1}{n-2}}$  for  $i = 1, \dots, M$ , then

$$\begin{aligned}r &\geq \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \inf_{\substack{V_1, \dots, V_N \geq 0, \\ l_1, \dots, l_N \geq 0, \\ V_1 + \dots + V_N = P_{\Gamma}(\tilde{\Gamma}), \\ l_1 + \dots + l_N \leq \mathcal{H}^1(\Sigma^*)}} \sum_{i=1}^M \left( V_i^{\frac{n}{n-1}} \right) + \sum_{i=M+1}^N V_i \left( \frac{V_i}{l_i} \right)^{\frac{1}{n-2}} \\ &= \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \inf_{\tilde{V}, \hat{V} \geq 0, M\tilde{V} + (N-M)\hat{V} = P_{\Gamma}(\tilde{\Gamma})} M\tilde{V}^{\frac{n}{n-1}} + [(N-M)\hat{V}]^{\frac{n-1}{n-2}} \mathcal{H}^1(\Sigma^*)^{\frac{-1}{n-2}},\end{aligned}$$

where we have used that the optimisation is convex in the  $V_i$  and  $l_i$  so that  $V_1 = \dots = V_M = \tilde{V}$ ,  $V_{M+1} = \dots = V_N = \hat{V}$ , and  $l_{M+1} = \dots = l_N = \frac{\mathcal{H}^1(\Sigma^*)}{N-M}$  are optimal. Substituting  $W = M\tilde{V}$  and replacing  $(N-M)\hat{V}$  by  $P_{\Gamma}(\tilde{\Gamma}) - W$ , the above estimate turns into

$$\begin{aligned}r &\geq \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \inf_{0 \leq W \leq P_{\Gamma}(\tilde{\Gamma}), \tilde{V} \geq \frac{W}{M}} W\tilde{V}^{\frac{1}{n-1}} + [P_{\Gamma}(\tilde{\Gamma}) - W]^{\frac{n-1}{n-2}} \mathcal{H}^1(\Sigma^*)^{\frac{-1}{n-2}} \\ &\geq \frac{C}{P_{\Gamma}(\tilde{\Gamma})} \inf_{0 \leq W \leq P_{\Gamma}(\tilde{\Gamma})} W^{\frac{n}{n-1}} N^{\frac{-1}{n-1}} + [P_{\Gamma}(\tilde{\Gamma}) - W]^{\frac{n-1}{n-2}} \mathcal{H}^1(\Sigma^*)^{\frac{-1}{n-2}} \\ &\geq \frac{C 2^{\frac{1-n}{n-2}}}{P_{\Gamma}(\tilde{\Gamma})} \min \left\{ P_{\Gamma}(\tilde{\Gamma})^{\frac{n}{n-1}} N^{\frac{-1}{n-1}}, P_{\Gamma}(\tilde{\Gamma})^{\frac{n-1}{n-2}} \mathcal{H}^1(\Sigma^*)^{\frac{-1}{n-2}} \right\},\end{aligned}$$

where the last step follows from distinguishing the two cases  $W > \frac{P_{\Gamma}(\tilde{\Gamma})}{2}$  and  $W \leq \frac{P_{\Gamma}(\tilde{\Gamma})}{2}$ . Inserting this estimate into the lower bound on  $\Delta\mathcal{E}^{\varepsilon,a,A}$  and using the fact  $N \leq \frac{\Delta\mathcal{E}^{\varepsilon,a,A}}{\varepsilon}$  and  $\mathcal{H}^1(\Sigma^*) \leq \frac{\Delta\mathcal{E}^{\varepsilon,a,A}}{\varepsilon}$ , we obtain

$$\begin{aligned}\Delta\mathcal{E}^{\varepsilon,a,A} &\geq C 2^{\frac{1-n}{n-2}} \sqrt{a^2 - 1} \min \left\{ P_{\Gamma}(\tilde{\Gamma})^{\frac{n}{n-1}} N^{\frac{-1}{n-1}}, P_{\Gamma}(\tilde{\Gamma})^{\frac{n-1}{n-2}} \mathcal{H}^1(\Sigma^*)^{\frac{-1}{n-2}} \right\} + (a - 1) P_{\Gamma}(\Gamma \setminus \tilde{\Gamma}) \\ &\geq C 2^{\frac{1-n}{n-2}} \sqrt{a^2 - 1} P_{\Gamma}(\tilde{\Gamma}) \min \left\{ \left( \varepsilon \frac{P_{\Gamma}(\tilde{\Gamma})}{\Delta\mathcal{E}^{\varepsilon,a,A}} \right)^{\frac{1}{n-1}}, \left( \varepsilon \frac{P_{\Gamma}(\tilde{\Gamma})}{\Delta\mathcal{E}^{\varepsilon,a,A}} \right)^{\frac{1}{n-2}} \right\} + (a - 1) P_{\Gamma}(\Gamma \setminus \tilde{\Gamma}).\end{aligned}$$

Now we distinguish three cases.

- If  $P_\Gamma(\tilde{\Gamma}) < \frac{1}{2}\mathcal{H}^{n-1}(A)$ , we have

$$\Delta\mathcal{E}^{\varepsilon,a,A} \geq (a-1)P_\Gamma(\Gamma \setminus \tilde{\Gamma}) \geq \frac{1}{2}\mathcal{H}^{n-1}(A)(a-1).$$

- If  $P_\Gamma(\tilde{\Gamma}) \geq \frac{1}{2}\mathcal{H}^{n-1}(A)$  and  $\varepsilon \frac{P_\Gamma(\tilde{\Gamma})}{\Delta\mathcal{E}^{\varepsilon,a,A}} < 1$ , we have

$$\Delta\mathcal{E}^{\varepsilon,a,A} \geq C2^{\frac{1-n}{n-2}}\sqrt{a^2-1}\frac{\mathcal{H}^{n-1}(A)}{2}\left(\varepsilon\frac{\mathcal{H}^{n-1}(A)}{2\Delta\mathcal{E}^{\varepsilon,a,A}}\right)^{\frac{1}{n-2}},$$

which can be solved to yield  $\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A)\sqrt{a^2-1}^{\frac{n-2}{n-1}}\varepsilon^{\frac{1}{n-1}}$ .

- If  $P_\Gamma(\tilde{\Gamma}) \geq \frac{1}{2}\mathcal{H}^{n-1}(A)$  and  $\varepsilon \frac{P_\Gamma(\tilde{\Gamma})}{\Delta\mathcal{E}^{\varepsilon,a,A}} \geq 1$ , we have

$$\Delta\mathcal{E}^{\varepsilon,a,A} \geq C2^{\frac{1-n}{n-2}}\sqrt{a^2-1}\frac{\mathcal{H}^{n-1}(A)}{2}\left(\varepsilon\frac{\mathcal{H}^{n-1}(A)}{2\Delta\mathcal{E}^{\varepsilon,a,A}}\right)^{\frac{1}{n-1}},$$

which can be solved to yield

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A)\sqrt{a^2-1}^{\frac{n-1}{n}}\varepsilon^{\frac{1}{n}} \geq \mathcal{H}^{n-1}(A)\min\left\{a-1, \sqrt{a^2-1}^{\frac{n-2}{n-1}}\varepsilon^{\frac{1}{n-1}}\right\}.$$

In summary,

$$\Delta\mathcal{E}^{\varepsilon,a,A} \gtrsim \mathcal{H}^{n-1}(A)\min\left\{a-1, \sqrt{a^2-1}^{\frac{n-2}{n-1}}\varepsilon^{\frac{1}{n-1}}\right\}.$$

□

### 4.3 Lower bound for $\mathcal{M}^{\varepsilon,A}$

To prove the lower bound of Theorem 2.4.2, it is actually easier to use a different non-dimensionalisation of the energy  $\mathcal{M}^{\varepsilon,m,A,L}$  introduced in Section 2.4. In fact, we shall prove a lower bound using the energy

$$\tilde{\mathcal{M}}^{\varepsilon,A} \equiv \mathcal{M}^{\varepsilon,m,A,1} \quad \text{with } m = \min\left\{1, \frac{1}{\mathcal{H}^{n-1}(A)}\right\}, \quad \tilde{\mathcal{M}}^{*,A} \equiv \inf_{\chi} \tilde{\mathcal{M}}^{0,A}[\chi] = \mathbf{m},$$

where  $\mathbf{m}$  denotes the total transported mass,

$$\mathbf{m} = \|\mu_0\|_{\text{fbm}} = m\mathcal{H}^{n-1}(A) = \min\{1, \mathcal{H}^{n-1}(A)\}.$$

Since optimal irrigation patterns  $\chi^*$  are known to be simple (that is, the patterns contain no loops, see [BCM09, Proposition 4.6]) and  $\mathbf{m} \leq 1$ , the maximum possible mass flux  $m_{\chi^*}(x)$  through any point  $x$  is no larger than one. This will allow us to bound the transport costs below by the costs for  $\varepsilon = 0$  since  $m_{\chi^*}(x) \leq 1$  implies  $s_{1-\varepsilon}^{\chi^*}(x) = m_{\chi^*}(x)^{-\varepsilon} \geq 1$  (see Definition 2.3.1). Indeed, let  $\chi^*$  be an optimal irrigation pattern between a given source  $\mu_+$  and sink  $\mu_-$  of equal mass no larger than one. For almost all  $p \in \Gamma$  we have

$$\int_I s_{1-\varepsilon}^{\chi^*}(\chi^*(p,t))|\dot{\chi}^*(p,t)| dt = \int_{\chi_p^*(I)} s_{1-\varepsilon}^{\chi^*}(x) d\mathcal{H}^1(x) \geq \int_{\chi_p^*(I)} d\mathcal{H}^1(x) \geq |\chi_p^*(1) - \chi_p^*(0)|.$$

Thus,

$$\begin{aligned} \mathcal{M}^{\varepsilon,\mu_+,\mu_-}[\chi^*] &= \int_\Gamma \int_I s_{1-\varepsilon}^{\chi^*}(\chi^*(p,t))|\dot{\chi}^*(p,t)| dt dP_\Gamma(p) \\ &\geq \int_\Gamma |\chi_p^*(1) - \chi_p^*(0)| dP_\Gamma(p) = \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y| d\mu(x,y) \geq W_1(\mu_+, \mu_-) \end{aligned}$$

for  $\mu = (\chi^*(\cdot, 0), \chi^*(\cdot, 1))_{\#} P_\Gamma$ , whose marginals are the irrigating and irrigated measures  $\mu_+$  and  $\mu_-$ .

Based on this relaxation result we will show the following.

**Theorem 4.3.1.** *There is a constant  $\tilde{C}_1 \equiv \tilde{C}_1(n)$  such that for  $\varepsilon < \tilde{C}_1$*

$$\min_{\chi} \tilde{\mathcal{M}}^{\varepsilon, A}[\chi] - \tilde{\mathcal{M}}^{*, A} \geq \tilde{C}_1 \mathfrak{m} \varepsilon \max\{|\log \varepsilon|, \log \mathcal{H}^{n-1}(A)\}.$$

If  $\mathcal{H}^{n-1}(A) \leq 1$ , then  $\tilde{\mathcal{M}}^{\varepsilon, A} = \mathcal{M}^{\varepsilon, A}$  so that the lower bound in Theorem 2.4.2 reduces to Theorem 4.3.1. If  $\mathcal{H}^{n-1}(A) > 1$ , then the relation

$$\mathcal{M}^{\varepsilon, A}[\chi] = \mathcal{H}^{n-1}(A)^{1-\varepsilon} \tilde{\mathcal{M}}^{\varepsilon, A}[\chi], \quad \mathcal{M}^{*, A} = \mathcal{H}^{n-1}(A) \tilde{\mathcal{M}}^{*, A},$$

immediately implies

$$\min_{\chi} \mathcal{M}^{\varepsilon, A}[\chi] - \mathcal{M}^{*, A} \geq \mathcal{H}^{n-1}(A) \left( \frac{1 + \tilde{C}_1 \varepsilon \max\{|\log \varepsilon|, \log \mathcal{H}^{n-1}(A)\}}{\mathcal{H}^{n-1}(A)^\varepsilon} - 1 \right),$$

from which the lower bound in Theorem 2.4.2 follows via  $\mathcal{H}^{n-1}(A)^{-\varepsilon} \geq 1 - \varepsilon \log \mathcal{H}^{n-1}(A)$ .

As usual, we would like to start by characterising a generic cross-section in terms of how many network pipes it intersects. However, unlike for the urban planning model, the maximum number of intersecting pipes cannot be deduced directly from the energy since the excess cost per pipe depends on the mass flux through it. Thus, we will for each flux separately treat the number of pipes with that flux. Analogously to before let  $\chi^*$  denote an optimal irrigation pattern of  $\tilde{\mathcal{M}}^{\varepsilon, A}$  with corresponding optimal network  $\Sigma^*$ , abbreviate

$$\Delta \tilde{\mathcal{M}}^{\varepsilon, A} = \tilde{\mathcal{M}}^{\varepsilon, A}[\chi^*] - \tilde{\mathcal{M}}^{*, A},$$

and recall that

$$t_p(s) = \min \{t \in I : \chi_p^*(t) \in \{x_n = s\}\}, \\ \xi^*(p, s) = \chi_p^*(t_p(s))$$

for  $s \in (0, 1)$ ,  $p \in \Gamma$  denote the time and position of particle  $p$  in cross-section  $\{x_n = s\}$ . For an easier notation, let us also introduce the corresponding solidarity classes and mass fluxes through  $x \in \mathbb{R}^n$ ,

$$[x]_{\xi^*} = \{q \in \Gamma : x \in \xi_q^*(I)\}, \\ m_{\xi^*}(x) = P_\Gamma([x]_{\xi^*}).$$

For each cross-section  $\{x_n = s\}$  we now introduce the number of intersecting pipes with mass flux  $c$  as well as a corresponding measure (discrete and with integer multiplicity) on the interval of possible mass fluxes,

$$\tilde{N}_s(c) = \mathcal{H}^0(\{x \in \{x_n = s\} : m_{\xi^*}(x) = c\}), \quad N_s = \tilde{N}_s \mathcal{H}^0 \llcorner (0, \mathfrak{m}].$$

It is known that  $\Sigma^* = \{x \in \mathbb{R}^n : m_{\chi^*}(x) > 0\}$  is rectifiable and represents a finite graph away from  $x_n = 0$  and  $x_n = 1$  (see [BCM09, Proposition 4.6] or [BCM05, Proposition 6.6]). Thus, for almost all  $s \in (0, 1)$ ,  $\Sigma^* \cap \{x_n = s\}$  is finite. Let us abbreviate

$$I_0 = \{s \in (0, 1) : \Sigma^* \cap \{x_n = s\} \text{ is finite}\}.$$

Clearly, for almost all  $s \in I_0$  the total flux through  $\{x_n = s\}$  is given by  $\int_{\{x_n = s\}} m_{\xi^*}(x) d\mathcal{H}^0(x) = P_\Gamma(\Gamma) = \|\mu_0\|_{\text{fbm}} = \mathfrak{m}$ , thus

$$\int_0^{\mathfrak{m}} c dN_s(c) = \mathfrak{m}.$$

The following lemma is the desired characterisation of a generic cross-section.

**Lemma 4.3.2.** *There is a generic  $s \in (0, 1)$  such that  $N_s(c)$  satisfies*

$$\int_0^{\mathfrak{m}} -c \log c dN_s(c) \leq \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\varepsilon}.$$

*Proof.* As mentioned above,  $\Sigma^* = \{x \in \mathbb{R}^n : m_{\chi^*}(x) > 0\}$  is rectifiable and represents a finite graph away from  $x_n = 0$  and  $x_n = 1$  so that we can calculate

$$\begin{aligned}\tilde{\mathcal{M}}^{\varepsilon,A}[\chi^*] &= \int_{\Gamma} \int_I s_{1-\varepsilon}^{\chi^*}(\chi^*(p,t)) |\dot{\chi}^*(p,t)| dt dP_{\Gamma} = \int_{\Gamma} \int_{\chi_p^*(I) \cap \Sigma^*} m_{\chi^*}(x)^{-\varepsilon} d\mathcal{H}^1(x) dP_{\Gamma} \\ &= \int_{\Sigma^*} \int_{\Gamma} m_{\chi^*}(x)^{-\varepsilon} \mathbf{1}_{\chi_p^*(I)}(x) dP_{\Gamma} d\mathcal{H}^1(x) = \int_{\Sigma^*} m_{\chi^*}(x)^{1-\varepsilon} d\mathcal{H}^1(x) \geq \int_{\Sigma^*} m_{\xi^*}(x)^{1-\varepsilon} d\mathcal{H}^1(x) \\ &\geq \int_{I_0} \int_{\Sigma^* \cap \{x_n=s\}} m_{\xi^*}(x)^{1-\varepsilon} d\mathcal{H}^0(x) ds = \int_{I_0} \int_0^{\mathfrak{m}} c^{1-\varepsilon} dN_s(c) ds.\end{aligned}$$

Assume by contradiction that  $\int_0^{\mathfrak{m}} -c \log c dN_s(c) > \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon,A}}{\varepsilon}$  for almost every  $s \in (0,1)$ . Then

$$\begin{aligned}\Delta \tilde{\mathcal{M}}^{\varepsilon,A} &= \tilde{\mathcal{M}}^{\varepsilon,A}[\chi^*] - \mathfrak{m} \geq \int_{I_0} \int_0^{\mathfrak{m}} c^{1-\varepsilon} dN_s(c) ds - \mathfrak{m} = \int_{I_0} \int_0^{\mathfrak{m}} (c^{1-\varepsilon} - c) dN_s(c) ds \\ &= \int_{I_0} \int_0^{\mathfrak{m}} c(c^{-\varepsilon} - 1) dN_s(c) ds \geq \int_{I_0} \int_0^{\mathfrak{m}} -c\varepsilon \log c dN_s(c) ds > \Delta \tilde{\mathcal{M}}^{\varepsilon,A},\end{aligned}$$

where we linearised in  $\varepsilon$  in the second last step. This yields a contradiction.  $\square$

To show Theorem 4.3.1, as earlier we bound the energy below by the transport costs for  $\varepsilon = 0$ , using the additional information from the previous lemma.

*Proof of Theorem 4.3.1.* Let  $\mu_s = \xi^*(\cdot, s)_{\#} P_{\Gamma}$  for the generic cross-section  $s$  from the previous lemma. By Remark 3.1.4 we have

$$\Delta \tilde{\mathcal{M}}^{\varepsilon,A} = \tilde{\mathcal{M}}^{\varepsilon,A}[\chi^*] - \tilde{\mathcal{M}}^{*,A} = \inf_{\chi} \mathcal{M}^{\varepsilon,A,\mu_0,\mu_s}[\chi] + \inf_{\chi} \mathcal{M}^{\varepsilon,A,\mu_s,\mu_1}[\chi] - \mathfrak{m}.$$

Now  $\mu_s$  is a linear combination of delta-distributions,

$$\mu_s = \sum_{\substack{c \in (0,\mathfrak{m}] \\ \tilde{N}_s(c) > 0}} \sum_{\substack{x \in \Sigma^* \cap \{x_n=s\} \\ m_{\xi^*}(x)=c}} c \delta_x =: \sum_{\substack{c \in (0,\mathfrak{m}] \\ \tilde{N}_s(c) > 0}} \nu_c.$$

Let  $\Gamma_c$  be the particles flowing through  $\nu_c$ , that is,  $\Gamma_c$  is defined by the relation  $\nu_c = \xi^*(\cdot, s)_{\#} P_{\Gamma} \llcorner \Gamma_c$ , and let  $\mu_{0,c}$  and  $\mu_{1,c}$  be initial and final distribution of these particles,

$$\mu_{0,c} = \xi^*(\cdot, 0)_{\#} P_{\Gamma} \llcorner \Gamma_c, \quad \mu_{1,c} = \xi^*(\cdot, 1)_{\#} P_{\Gamma} \llcorner \Gamma_c.$$

Lemma 4.1.1 together with  $\sum_{c \in (0,\mathfrak{m}], \tilde{N}_s(c) > 0} c \tilde{N}_s(c) = \mathfrak{m}$  then implies

$$\begin{aligned}\inf_{\chi} \mathcal{M}^{\varepsilon,A,\mu_0,\mu_s}[\chi] + \inf_{\chi} \mathcal{M}^{\varepsilon,A,\mu_s,\mu_1}[\chi] &\geq \sum_{\substack{c \in (0,\mathfrak{m}] \\ \tilde{N}_s(c) > 0}} W_1(\mu_{0,c}, \nu_c) + W_1(\nu_c, \mu_{1,c}) \\ &\geq \sum_{\substack{c \in (0,\mathfrak{m}] \\ \tilde{N}_s(c) > 0}} c \tilde{N}_s(c) \left[ s + C \min \left\{ \left( \frac{c}{m\omega_{n-1}} \right)^{\frac{2}{n-1}} / s, \left( \frac{c}{m\omega_{n-1}} \right)^{\frac{1}{n-1}} \right\} \right] \\ &\quad + \sum_{\substack{c \in (0,\mathfrak{m}] \\ \tilde{N}_s(c) > 0}} c \tilde{N}_s(c) \left[ 1 - s + C \min \left\{ \left( \frac{c}{m\omega_{n-1}} \right)^{\frac{2}{n-1}} / (1-s), \left( \frac{c}{m\omega_{n-1}} \right)^{\frac{1}{n-1}} \right\} \right] \\ &\geq \mathfrak{m} + \int_0^{\mathfrak{m}} c \tilde{C} \min \left\{ \left( \frac{c}{m} \right)^{\frac{2}{n-1}}, \left( \frac{c}{m} \right)^{\frac{1}{n-1}} \right\} dN_s(c)\end{aligned}$$

for some constant  $\tilde{C} \equiv \tilde{C}(n) > 0$ . Consequently, exploiting  $m, \mathfrak{m} \leq 1$ ,

$$\begin{aligned}\Delta \tilde{\mathcal{M}}^{\varepsilon,A} &\geq \tilde{C} \int_0^{\mathfrak{m}} c \min \left\{ \left( \frac{c}{m} \right)^{\frac{2}{n-1}}, \left( \frac{c}{m} \right)^{\frac{1}{n-1}} \right\} dN_s(c) \\ &\geq \frac{\tilde{C}}{m^{\frac{1}{n-1}}} \int_0^{\mathfrak{m}} c \min \left\{ c^{\frac{2}{n-1}}, c^{\frac{1}{n-1}} \right\} dN_s(c) \geq \frac{\tilde{C}}{m^{\frac{1}{n-1}}} \int_0^{\mathfrak{m}} c^{\frac{n+1}{n-1}} dN_s(c).\end{aligned}$$

Together with the previous lemma and  $\int_0^{\mathbf{m}} c \, dN_s(c) = \mathbf{m}$  this implies

$$\Delta \tilde{\mathcal{M}}^{\varepsilon, A} \geq \frac{\tilde{C}}{m^{\frac{1}{n-1}}} \inf_{\substack{N \in \text{fbm}([0, \mathbf{m}]) \\ \int_0^{\mathbf{m}} c \, dN(c) = \mathbf{m} \\ \int_0^{\mathbf{m}} -c \log c \, dN(c) \leq \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\varepsilon}}} \int_0^{\mathbf{m}} c^{\frac{n+1}{n-1}} \, dN(c),$$

where the right-hand side represents a linear program. The next lemma now implies

$$\Delta \tilde{\mathcal{M}}^{\varepsilon, A} \geq \frac{\tilde{C}}{m^{\frac{1}{n-1}}} \mathbf{m} \exp \left( -\frac{2\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{(n-1)\mathbf{m}\varepsilon} \right)$$

and thus  $-\log \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\tilde{C}\mathbf{m}} - \frac{\log m}{n-1} \leq \frac{\tilde{C}}{n-1} \frac{1}{\varepsilon} \frac{2\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\tilde{C}\mathbf{m}}$ . From the upper bound we already know  $\Delta \tilde{\mathcal{M}}^{\varepsilon, A} \leq C_2 \mathbf{m} \varepsilon |\log \varepsilon|$  for a constant  $C_2 \equiv C_2(n) > 0$ , hence  $\log \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\tilde{C}\mathbf{m}} \leq 0$  for  $\varepsilon$  small enough such that  $\varepsilon |\log \varepsilon| \leq \frac{\tilde{C}}{C_2}$ . Thus, the above inequality implies  $\Delta \tilde{\mathcal{M}}^{\varepsilon, A} \geq \frac{1}{2} \mathbf{m} \varepsilon |\log m|$  as well as  $\left| \log \frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\tilde{C}\mathbf{m}} \right| \leq \frac{\tilde{C}}{n-1} \frac{1}{\varepsilon} \frac{2\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\tilde{C}\mathbf{m}}$  from which we infer  $\frac{\Delta \tilde{\mathcal{M}}^{\varepsilon, A}}{\mathbf{m}} \geq \tilde{C}_1 \varepsilon |\log \varepsilon|$  for a constant  $\tilde{C}_1 \equiv \tilde{C}_1(n)$ .  $\square$

Obviously, the proof of Theorem 4.3.1 reduces the energy estimate to a convex optimisation problem. The role of the following lemma is simply to provide the necessary lower bound on that problem via convex duality.

**Lemma 4.3.3.** *For any  $\mathbf{m}, D > 0$  we have*

$$\inf_{\substack{N \in \text{fbm}([0, \mathbf{m}]) \\ \int_0^{\mathbf{m}} c \, dN(c) = \mathbf{m} \\ \int_0^{\mathbf{m}} -c \log c \, dN(c) \leq D}} \int_0^{\mathbf{m}} c^{\frac{n+1}{n-1}} \, dN(c) \geq \mathbf{m} \exp \left( -\frac{2D}{(n-1)\mathbf{m}} \right).$$

*Proof.* The optimisation problem is equivalent to finding

$$\rho := \inf_{N \in \text{rca}([0, \mathbf{m}])} G^*(N) + H^*(N)$$

for the Legendre–Fenchel duals

$$\begin{aligned} G^* : \text{rca}([0, \mathbf{m}]) &\rightarrow \mathbb{R}, \quad G^*(N) = \int_0^{\mathbf{m}} c^{\frac{n+1}{n-1}} \, dN + I_{\{N \geq 0\}}(N) \\ H^* : \text{rca}([0, \mathbf{m}]) &\rightarrow \mathbb{R}, \quad H^*(N) = I_{\{\int_0^{\mathbf{m}} c \, dN = \mathbf{m}\}}(N) + I_{\{\int_0^{\mathbf{m}} -c \log c \, dN \leq D\}}(N) \end{aligned}$$

of

$$\begin{aligned} G : \mathcal{C}([0, \mathbf{m}]) &\rightarrow \mathbb{R}, \quad G(\phi) = I_{\{\phi \leq 0\}} \left( c \mapsto \phi(c) - c^{\frac{n+1}{n-1}} \right) \\ H : \mathcal{C}([0, \mathbf{m}]) &\rightarrow \mathbb{R}, \quad H(\phi) = \begin{cases} \lambda \mathbf{m} - \kappa D & \text{if } \phi(c) = \lambda c + \kappa c \log c \text{ for } \lambda \in \mathbb{R}, \kappa \leq 0, \\ \infty & \text{else,} \end{cases} \end{aligned}$$

where  $\mathcal{C}$  denotes the set of continuous functions,  $I_S$  denotes the indicator function of a set  $S$ , that is,  $I_S(x) = 0$  if  $x \in S$  and  $I_S(x) = \infty$  else. By Fenchel–Rockafellar duality,

$$\rho \geq \sup_{\phi \in \mathcal{C}([0, \mathbf{m}])} -G(-\phi) - H(\phi) = \sup_{\substack{\lambda \in \mathbb{R}, \kappa \leq 0, \\ \lambda c + \kappa c \log c \geq -c^{\frac{n+1}{n-1}}}} -\lambda \mathbf{m} + \kappa D = \sup_{\substack{\lambda \in \mathbb{R}, \kappa \leq 0, \\ \lambda - \kappa \log c \leq c^{\frac{2}{n-1}}}} \lambda \mathbf{m} + \kappa D.$$

Any admissible choice of  $\lambda, \kappa$  yields a lower bound. We choose  $\lambda, \kappa$  such that  $\lambda - \kappa \log c$  actually touches the curve  $c^{\frac{2}{n-1}}$  in a point  $c^*$  tangentially, that is, the function values and derivatives of both curves coincide,

$$\lambda - \kappa \log c^* = (c^*)^{\frac{2}{n-1}}, \quad -\frac{\kappa}{c^*} = \frac{2}{n-1} (c^*)^{\frac{3-n}{n-1}}.$$

In particular,  $\kappa = -\frac{2}{n-1}(c^*)^{\frac{2}{n-1}}$  and  $\lambda = (c^*)^{\frac{2}{n-1}}(1 - \frac{2}{n-1} \log c^*)$ . A maximisation of  $\lambda \mathfrak{m} + \kappa D$  for  $c^*$  now yields  $c^* = \exp(-\frac{D}{\mathfrak{m}})$  and thus

$$\lambda = \exp\left(-\frac{2D}{\mathfrak{m}(n-1)}\right) \left(\frac{2D}{\mathfrak{m}(n-1)} + 1\right), \quad \kappa = -\frac{2}{n-1} \exp\left(-\frac{2D}{\mathfrak{m}(n-1)}\right),$$

which produces

$$\rho \geq \lambda \mathfrak{m} + \kappa D = \mathfrak{m} \exp\left(-\frac{2D}{\mathfrak{m}(n-1)}\right).$$

□

*Remark 4.3.4.* The test function  $N(c) = \frac{\mathfrak{m}}{c^*} \delta(c - c^*)$  for a properly chosen  $c^*$  actually shows that the inequality in Lemma 4.3.3 is an equality.

## 5 Discussion

In the following we briefly discuss different geometric settings as well as the relation of network optimisation to other pattern formation problems.

### 5.1 Geometry variations

It is out of the scope of this paper to rigorously prove the scaling laws also for variations of our geometric setting, however, we would like to provide a little discussion of the dependence on geometry. First of all notice that all lower bound proofs are still valid if the hypersquare  $A$  is replaced by an arbitrary subset of  $\mathbb{R}^{n-1}$ . Likewise, the upper bound constructions can easily be adapted to non-square geometries of  $A$ , one only has to take care that the width of the tree-like structures does not extend beyond  $A$  and that the construction has to be adapted to properly follow  $\partial A$ . It seems also rather clear that the analysis can be adapted to a non-uniform mass distribution on  $A \times \{0\}$  and  $A \times \{L\}$  as long as the density is uniformly bounded and strictly bounded away from 0. Another geometric setting which may be relevant for various applications (thinking for instance of plants or blood vessel systems) is where  $\mu_0$  is replaced by a Dirac measure,

$$\mu_0 = \mathcal{H}^{n-1}(A)\delta_0, \quad \mu_1 = \mathcal{H}^{n-1}\llcorner(A \times \{1\}).$$

Here, the upper bound construction can be obtained in a similar manner as for our usual geometry: One can choose elementary cells which are mapped into the standard rectangular elementary cells after applying the transformation

$$(x_1, \dots, x_n) \mapsto \left(\frac{x_1}{x_n}, \dots, \frac{x_{n-1}}{x_n}, x_n\right),$$

which is affine on each cross-section  $\{x_n = t\}$  and which maps the cone  $\{t(A \times \{1\}) : t \in [0, 1]\}$  onto the rectangle  $A \times [0, 1]$ . Such constructions provide the same energy scaling as for the geometry actually considered in this article. For the lower bound, for instance on the branched transport model, we may again consider the mass flux  $\mu_t$  on a generic cross-section  $t$  and then bound the energy from below by the transport costs from  $\mu_0$  to  $\mu_t$ ,  $\inf_{\chi} \mathcal{M}^{\varepsilon, \mu_0, \mu_t}[\chi] \geq W_1(\mu_0, \mu_t)$ , and from  $\mu_t$  to  $\mu_1$ ,  $\inf_{\chi} \mathcal{M}^{\varepsilon, \mu_t, \mu_1}[\chi] \geq W_1(\mu_t, \mu_1)$ . Now, however, it is no longer possible to separately estimate  $W_1(\mu_0, \mu_t)$  and  $W_1(\mu_t, \mu_1)$ . Instead, a variation of Lemma 4.1.1 yields the necessary bound on  $W_1(\mu_0, \mu_t) + W_1(\mu_t, \mu_1) - W_1(\mu_0, \mu_1)$ , resulting in the same energy scaling.

### 5.2 Properties characteristic for network optimisation

Here we would like to briefly recapitulate the observed phenomena and relate network optimisation to other pattern formation problems. Both urban planning and branched transport lead to branched pipe networks. However, one can observe a difference between both models: While for branched transport the excess energy over the limit case  $\varepsilon = 0$  is more or less evenly distributed across all branching levels, this is true for urban planning only in three dimensions. In two dimensions, the major contribution to the urban planning excess energy stems from the coarsest network structures, while in dimensions higher than three the dominant energetic effect stems from the even distribution of mass at the boundary.

The urban planning model is closer in structure to classical pattern formation models (such as martensite–austenite transformations [KM92] or intermediate states in type-I superconductors [CKO08])

than branched transport. Indeed, similarly to urban planning, the interfaces in many classical pattern formation models are directly penalised with cost  $\varepsilon$ , while in branched transport the preference for shorter and fewer pipes is rather implicitly encoded in the mass flux cost. The urban planning model is particularly related to patterns observed in intermediate states in type-I superconductors [CCKO08] or compliance minimisation [KW14] in which also a structure is sought that has to conduct a flux (in one case a magnetic, in the other a force flux). Urban planning is special, though, in that the conducting structures are really one-dimensional, while in the other mentioned problems they exhibit a finite width. Intuitively, this also implies that the other problems have more design freedom so that their analysis should be a little more complicated. And indeed, in our lower bound proofs it was relatively straightforward to explicitly characterise cross-sections through the network, which is much more difficult in problems without one-dimensional structures.

## 6 Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Cells-in-Motion Cluster of Excellence (EXC 1003-CiM), University of Münster, Germany. B.W.'s research was supported by the Alfried Krupp Prize for Young University Teachers awarded by the Alfried Krupp von Bohlen und Halbach-Stiftung.

## References

- [BB05] Alessio Brancolini and Giuseppe Buttazzo. Optimal networks for mass transportation problems. *ESAIM Control Optim. Calc. Var.*, 11(1):88–101 (electronic), 2005.
- [BCM05] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. Traffic plans. *Publ. Mat.*, 49(2):417–451, 2005.
- [BCM08] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. The structure of branched transportation networks. *Calc. Var. Partial Differential Equations*, 32(3):279–317, 2008.
- [BCM09] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. *Optimal transportation networks*, volume 1955 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Models and theory.
- [BG15] Peter Bella and Michael Goldman. Nucleation barriers at corners for a cubic-to-tetragonal phase transformation. *Proc. R. Soc. Edinb., Sect. A, Math.*, 145(4):715–724, 2015.
- [BK14] Peter Bella and Robert V. Kohn. Wrinkles as the result of compressive stresses in an annular thin film. *Comm. Pure Appl. Math.*, 67(5):693–747, 2014.
- [BK15] Jacob Bedrossian and Robert V. Kohn. Blister patterns and energy minimization in compressed thin films on compliant substrates. *Comm. Pure Appl. Math.*, 68(3):472–510, 2015.
- [BOS02] Giuseppe Buttazzo, Edouard Oudet, and Eugene Stepanov. Optimal transportation problems with free Dirichlet regions. In *Variational methods for discontinuous structures*, volume 51 of *Progr. Nonlinear Differential Equations Appl.*, pages 41–65. Birkhäuser, Basel, 2002.
- [BPSS09] Giuseppe Buttazzo, Aldo Pratelli, Sergio Solimini, and Eugene Stepanov. *Optimal urban networks via mass transportation*, volume 1961 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.
- [BS14] Alessio Brancolini and Sergio Solimini. Fractal regularity results on optimal irrigation patterns. *J. Math. Pures Appl. (9)*, 102(5):854–890, 2014.
- [BW15] Alessio Brancolini and Benedikt Wirth. Equivalent formulations for the branched transport and urban planning problems. *ArXiv e-prints*, 1509.06698, September 2015.

- [CC15] Allan Chan and Sergio Conti. Energy scaling and branched microstructures in a model for shape-memory alloys with  $SO(2)$  invariance. *Math. Models Methods Appl. Sci.*, 25(6):1091–1124, 2015.
- [CCKO08] Rustum Choksi, Sergio Conti, Robert V. Kohn, and Felix Otto. Ground state energy scaling laws during the onset and destruction of the intermediate state in a type I superconductor. *Comm. Pure Appl. Math.*, 61(5):595–626, 2008.
- [CKO99] Rustum Choksi, Robert V. Kohn, and Felix Otto. Domain branching in uniaxial ferromagnets: a scaling law for the minimum energy. *Comm. Math. Phys.*, 201(1):61–79, 1999.
- [CKO04] R. Choksi, R. V. Kohn, and F. Otto. Energy minimization and flux domain structure in the intermediate state of a type-I superconductor. *J. Nonlinear Sci.*, 14(2):119–171, 2004.
- [Con] Sergio Conti. Personal communication.
- [GZ14] M. Goldman and B. Zwicknagl. Scaling law and reduced models for epitaxially strained crystalline films. *SIAM J. Math. Anal.*, 46(1):1–24, 2014.
- [KKO13] Hans Knüpfer, Robert V. Kohn, and Felix Otto. Nucleation barriers for the cubic-to-tetragonal phase transformation. *Comm. Pure Appl. Math.*, 66(6):867–904, 2013.
- [KM92] R. V. Kohn and S. Müller. Branching of twins near an austenite-twinned-martensite interface. *Philosophical Magazine A*, 1992.
- [KM94] Robert V. Kohn and Stefan Müller. Surface energy and microstructure in coherent phase transitions. *Comm. Pure Appl. Math.*, 47(4):405–435, 1994.
- [KW14] Robert V. Kohn and Benedikt Wirth. Optimal fine-scale structures in compliance minimization for a uniaxial load. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 470(2170), 2014.
- [KW15] Robert Kohn and Benedikt Wirth. Optimal fine-scale structures in compliance minimization for a shear load. *Comm. Pure Appl. Math.*, 2015.
- [MS09] Francesco Maddalena and Sergio Solimini. Transport distances and irrigation models. *J. Convex Anal.*, 16(1):121–152, 2009.
- [MS13] Francesco Maddalena and Sergio Solimini. Synchronic and asynchronous descriptions of irrigation problems. *Adv. Nonlinear Stud.*, 13(3):583–623, 2013.
- [MSM03] F. Maddalena, S. Solimini, and J.-M. Morel. A variational model of irrigation patterns. *Interfaces Free Bound.*, 5(4):391–415, 2003.
- [Roy88] Halsey L. Royden. *Real analysis*. Macmillan Publishing Company, New York, third edition, 1988.
- [Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.